

STIC-Biotech/ChemLib

72632

From: Rao, Manjunath N.  
Sent: Tuesday, August 06, 2002 3:09 PM  
To: STIC-Biotech/ChemLib  
Subj ct: Sequence search request for 09/742,690

From: Manjunath N. Rao  
Art Unit 1652, Room 10A11  
Mail Box in Room 10C 01  
Phone: 306-5681

Date: 8-6-02

Please search the following as soon as possible for application with serial number **09/742,690**

**SEQ ID NO: 1** against all commercial nucleic acid databases including issued patents database and pending application database and provide a **print of all results**.

**SEQ ID NO: 1** against all commercial protein databases including issued patents database and pending application database and provide a **print of all results**.

If you have any questions please call me at the above phone number.

Thanks

Manjunath N. Rao, Ph.D.  
Biotechnology Patent Examiner  
Art Unit 1652, Room 10A11  
Crystal Mall 1, USPTO.

STIC-Biotech/ChemLib  
Aug 6 2002  
10:00 AM

Point of Contact  
P. Sheppard  
Searcher: Telephone number: (703) 308-4499  
Phone: \_\_\_\_\_  
Location: \_\_\_\_\_  
Date Picked Up: \_\_\_\_\_  
Date Completed: 8/14/02  
Searcher Prep/Review: \_\_\_\_\_  
Clerical: \_\_\_\_\_  
Online time: \_\_\_\_\_

TYPE OF SEARCH:  
NA Sequences: \_\_\_\_\_  
AA Sequences: \_\_\_\_\_  
Structures: \_\_\_\_\_  
Bibliographic: \_\_\_\_\_  
Litigation: \_\_\_\_\_  
Full text: \_\_\_\_\_  
Patent Family: \_\_\_\_\_  
Other: \_\_\_\_\_

VENDOR/COST (where applic.)  
STN: \_\_\_\_\_  
DIALOG: \_\_\_\_\_  
Questel/Orbit: \_\_\_\_\_  
DRLink: \_\_\_\_\_  
Lexis/Nexis: \_\_\_\_\_  
Sequence Sys.: \_\_\_\_\_  
WWW/Internet: \_\_\_\_\_  
Other (specify): \_\_\_\_\_

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 13, 2002, 06:44:47 ; Search time 1772.03 Seconds  
(without alignments)  
4250.098 Million cell updates/sec

Title: US-09-742-690-1  
Perfect score: 558  
Sequence: 1 tcgagaaaagacaggtgcag.....ctcccagtgctgtgtaataag 558

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST:\*

1:	em_estba:*
2:	em_esthum:*
3:	em_estin:*
4:	em_estmu:*
5:	em_estov:*
6:	em_estpl:*
7:	em_estro:*
8:	em_htc:*
9:	gb_est1:*
10:	gb_est2:*
11:	gb_htc:*
12:	gb_gss:*
13:	em_gss_hum:*
14:	em_gss_inv:*
15:	em_gss_pln:*
16:	em_gss_vrt:*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	199.2	35.7	1071	10	BM454450
2	195.2	35.0	663	10	AGENCOURT
3	194.4	34.8	1164	10	BF974771
4	191.8	34.4	669	10	BF974771
5	189.2	33.9	687	10	BF974771
6	189.2	33.9	832	10	BF974771
7	187.6	33.6	518	9	AW402029
8	185.8	33.3	802	10	BF974771
9	185.2	33.2	571	10	BF974771
10	182.8	32.8	648	10	BF974771
11	182.4	32.7	559	9	AW401799
12	182.4	32.7	759	10	BF974771
13	182.2	32.7	493	9	AW402069
14	182	32.6	471	9	AW403220
15	182	32.6	516	9	AW401728
16	181.4	32.5	800	10	BM007901
17	180.6	32.4	874	10	BI755805

18	180.4	32.3	445	9	AW408316	AW408316 UI-HF-BK0
19	179.6	32.2	846	10	BF974771	BF974771 UI-HF-BK0
20	179.2	32.1	705	10	BF974771	BF974771 UI-HF-BK0
21	178.8	32.0	750	10	BF974771	BF974771 UI-HF-BK0
22	178.6	32.0	503	9	AW402572	AW402572 UI-HF-BK0
23	178.2	31.9	735	10	BF974771	BF974771 UI-HF-BK0
24	177.8	31.9	487	9	AW403862	AW403862 UI-HF-BK0
25	177.6	31.8	695	10	BF974771	BF974771 UI-HF-BK0
26	177.2	31.8	870	10	BF974771	BF974771 UI-HF-BK0
27	177.2	31.8	908	10	BF974771	BF974771 UI-HF-BK0
28	177	31.7	396	9	AW407843	AW407843 UI-HF-BK0
29	177	31.7	912	10	BF974771	BF974771 UI-HF-BK0
30	176.2	31.6	930	10	BF974771	BF974771 UI-HF-BK0
31	175.8	31.5	717	10	BF974771	BF974771 UI-HF-BK0
32	175.6	31.5	412	9	AW402895	AW402895 UI-HF-BK0
33	175.6	31.5	595	10	BF974771	BF974771 UI-HF-BK0
34	174.8	31.3	940	10	BF974771	BF974771 UI-HF-BK0
35	174.6	31.3	740	10	BF974771	BF974771 UI-HF-BK0
36	174.6	31.3	741	10	BF974771	BF974771 UI-HF-BK0
37	174.4	31.3	511	9	AW402613	AW402613 UI-HF-BK0
38	174.4	31.3	678	10	BF974771	BF974771 UI-HF-BK0
39	174.4	31.3	870	10	BF974771	BF974771 UI-HF-BK0
40	174.4	31.3	991	10	BF974771	BF974771 UI-HF-BK0
41	174.2	31.2	536	9	AW402624	AW402624 UI-HF-BK0
42	174.2	31.2	769	10	BM007986	BM007986 603617724
43	174	31.2	491	9	AW402907	AW402907 UI-HF-BK0
44	174	31.2	704	10	BF974771	BF974771 UI-HF-BK0
45	173.8	31.1	483	9	AW403901	AW403901 UI-HF-BK0

ALIGNMENTS

RESULT 1  
BM454450  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT  
FEATURES  
Source

BM454450  
AGENCOURT\_6406467 NIH\_MGC\_92 Homo sapiens cDNA clone IMAGE:5582873  
5', mRNA sequence.  
BM454450  
EST.  
human.  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 1071)  
NIH-MGC <http://mgc.nci.nih.gov/>.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs-r@mail.nih.gov](mailto:cgapbs-r@mail.nih.gov)  
Tissue Procurement: ATCC  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: LLAM12345 row: a column: 18  
High quality sequence stop: 425.  
Location/Qualifiers  
1..1071  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5582873"  
/clone\_lib="NIH\_MGC\_92"  
/tissue\_type="embryonal carcinoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: testis; Vector: pCMV-SPORT6; Site: 1; NotI; Site 2; Sali; Cloned unidirectionally; oligo-dT primed. Average insert size 2.5 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH\_MGC Library."

```
BASE COUNT      195 a      304 c      303 g      269 t
ORIGIN

Query Match      35.7%; Score 199.2; DB 10; Length 1071;
Best Local Similarity 67.9%; Pred. No. 2.2e-48;
Matches 326; Conservative 0; Mismatches 143; Indels 11; Gaps 3;

QY 13 aggtgcagctgcaggagtcagggggaggttggtgcaggctggggctctctgagactct 72
|||||
Db 221 AGGTGCAGCTGGCGGAGTCTGGGGGAGGTGTGGTACGCGCTGGGGGTCCCTGAGACTCT 280
|||||

QY 73 cctgtgcagcctcgggacgcgccaccagtggtcatggtcactatggtatgggctggttcc 132
|||||
Db 281 CCTTTCAGCCTCTGGATTACCC-----TTTGATGATTATGGCATGAGCTGGGTCC 331
|||||

QY 133 gccagggtccagggaagcagcgtgagtttgcgcagctattagggtggagtggtaagaga 192
|||||
Db 332 GCCAAGCTCCAGGGAAGGGGCTGGAGTGGGTCTCTGGTATTAAATGGAATGGTGTAGCA 391
|||||

QY 193 catggtataaagactccgtgaaggccgattccaccatctccagagataaacgccaaagacta 252
|||||
Db 392 CAGGTTATGCAGACTCTGTGAAGGGCCGATTCAACCATTCTCCAGAGACAACGCCAAGAACT 451
|||||

QY 253 cgggtttatctgcaaatgaacagcctgaaacccctgaagatacagcgccgtttattattgtgc-c 311
|||||
Db 452 CCCTGTATCTGCAAAATGAACAGCCTGAGAGTCGAGAGTGGAGACACGGCCCTTGATCACTGTGCGA 511
|||||

QY 312 gctcgaccggtccgctggtgatgatatttccctgcgggttggtttgactactcgggcccag 371
|||||
Db 512 GAGATTACAGTACGTTACTAATGGTGTATGCTATACGGAGTGGTTCGACCCCTGGGGCCAG 571
|||||

QY 372 gggaccaggtcacccgtctctccaggt-atctcatcacatcaccatcacggatccacctc 430
|||||
Db 572 GGAACCCCTGGTCACCGTCTCTCCAGGCGAGTCTCACACCCCTCTCTGAGTCCACTTA 631
|||||

QY 431 cattgaaggtcgtaccagctcactacgtcaggtcaggtggtggtgattggttactccggtcc 490
|||||
Db 632 GGGAGACTCCCCCTTGCCAGGGTCTCAGGGTTAACAGTCTCTGGAGGCCATTTTGGAGGTCC 691
|||||

RESULT 2
BG686759      663 bp      mRNA      linear      EST 01-MAY-2001
LOCUS
DEFINITION 602650729F1 NIH_MGC_48 Homo sapiens cDNA clone IMAGE:4763215 5',
mRNA sequence.
ACCESSION BG686759
VERSION BG686759.1 GI:13918156
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE 1 (bases 1 to 663)
JOURNAL NIH-MGC http://mgc.nci.nih.gov/.
COMMENT National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-re@mail.nih.gov
Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM1618 row: i column: 08
High quality sequence stop: 659.
Location/Qualifiers
1. .663
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4763215"
```

```
/clone_lib="NIH_MGC_48"
/tissue_type="primary B-cells from tonsils (cell line)"
/lab_host="DH10B (phage-resistant)"
/note="Organ: B-cells; Vector: pOTB7; Site_1: XhoI;
Site_2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Size-selected >500bp
for average insert size 1.8kb. Library constructed by Ling
Hong in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

BASE COUNT      138 a      185 c      186 g      154 t
ORIGIN

Query Match      35.0%; Score 195.2; DB 10; Length 663;
Best Local Similarity 71.1%; Pred. No. 2.7e-47;
Matches 276; Conservative 0; Mismatches 103; Indels 9; Gaps 1;

QY 10 gacaggtgcagctgcaggatcagggggaggattggtgcaggctggggctctctgagac 69
|||||
Db 126 GTCAGGTGCAGCTGGTGGAGTCTGGGGGAGGCGTGGTCCAGCCTGGGAGGTCCCTGAGAC 185
|||||

QY 70 tctcctgtgcagcctcgggacgcgccaccagtggtcatggtcactatggtatgggctggt 129
|||||
Db 186 TCTCCTGTGCAGCCTCTGGATTCACTTTCAGTAG-----CTATGCTATGCACACTGGG 236
|||||

QY 130 tccgccaggttccagggaaggagcgtgagtttgcgcagctattagggtggagtgtaaaag 189
|||||
Db 237 TCCGCCAGGCTCCAGGCAAGGGGCTGGAGTGGGTGGCAGTTATATCATATGATGGAAGCA 296
|||||

QY 190 agacatggtataaagactccgtgaaggcccgattcaccatctccagagataaacgccaaga 249
|||||
Db 297 ATAAATACTACGCAGACTCCCTGAAGGGCCGATTCACCATCTCCAGAGACAATTTCCAAGA 356
|||||

QY 250 ctacgggtttatctgcaaatgaacagcctgaaacctgaagatacagccgtttattattgtg 309
|||||
Db 357 ACACGCTGTATCTGCAAAATGAACAGCCTGAGAGCTGAGGACACGCGCTGTGTATTACTGTG 416
|||||

QY 310 ccgctcgaccggtccgctggtgatgatatttccctgcgggttggtttgactactcggggccc 369
|||||
Db 417 CGAGAGATCTTTTTCGTATAGCAGTGGCTGGCCATCGGGTACCTCGGCTACTGGGGGCC 476
|||||

QY 370 aggggaccaggtcacccgtctcctcagg 397
|||||
Db 477 AGGGAACCCCTGGTCAACCGTCTCCTCAGG 504
|||||

RESULT 3
BF974771      1164 bp      mRNA      linear      EST 22-JAN-2001
LOCUS
DEFINITION 602245420F1 NIH_MGC_48 Homo sapiens cDNA clone IMAGE:4336541 5',
mRNA sequence.
ACCESSION BF974771
VERSION BF974771.1 GI:12341986
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE 1 (bases 1 to 1164)
JOURNAL NIH-MGC http://mgc.nci.nih.gov/.
COMMENT National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-re@mail.nih.gov
Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
```







```

IMAGE:3054247 5', mRNA sequence.
ACCESSION AW402029
VERSION AW402029.1 GI:6920715
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 518)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Eco RI site shown at the beginning of the sequence.
Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
cDNA Library Preparation: M.B. Soares Lab
CDNA Library Arrayed by: M.B. Soares Lab
DNA Sequencing by: M.B. Soares Lab
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Seq primer: M13 Forward.
FEATURES
    source
        1..518
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /clone="IMAGE:3054247"
            /clone_lib="NIH_MGC_36"
            /tissue_type="lymph"
            /cell_type="germinal center B cells"
            /cell_line="MGC85"
            /lab_host="DH10B (LTI)"
            /note="Vector: pT7T3-Pac; Site_1: NotI; Site_2: Eco RI;
            Constructed from size fractionated cytoplasmic mRNA
            (0.5-1.5kb). Directionally cloned. Cells provided by Louis
            M. Staudt, Ph.D. Library preparation by Maria de Fatima
            Bonaldo, Ph.D. and M. Bento Soares, Ph.D."
BASE COUNT 99 a 149 c 148 g 121 t 1 others
ORIGIN
Query Match 33.6%; Score 187.6; DB 9; Length 518;
Best Local Similarity 73.7%; Pred. No. 4.5e-45;
Matches 286; Conservative 0; Mismatches 84; Indels 18; Gaps 3;
QY 13 aggtgcagctgcaggagtcagggaggagtggtgcaggctggggctctctgagactct 72
Db 54 ACGTGCAGCTGGTGGAGTCTGGGGGAGGCTTGGTACAGCCTGGGGGGTCCCTGAGACTCT 113
QY 73 cctgtgcagcctcgggacgcgcaccagtggtcatggtcactatggtatgggctggtcc 132
Db 114 CTTGTGCAGCCTCTGGATTACCTTTAGCAG-----CTATGCCATGACCTGGGTCC 164
QY 133 gccaggttccagggaagagcgtgagtttgcgcagctattaggtggagtggttaaagaga 192
Db 165 GCCAGGCTCCAGGGAAGGGCTGGAGTGGGTCTCAATTATTAGTGTGTACTAGCA 224
QY 193 catggtataaagactccgtgaaggccgattcaccatccagagataaacgccaaagata 252
Db 225 CATACTACGCAGACTCCGTGAAGGGCCGGTTTCATCATCTCCAGAGACAATTCCAAGAACA 284
QY 253 cggtttatctgcaaatgaacagcctgaaacctgaagatacggccgtttattattgtgccg 312
Db 285 CGCTGTTTCTGCAAAATGAACAGCCTGAGAGTCGAGGACACGCCGTATATTACTGTGC-- 342
QY 313 ctcgaccggtccgcgtgatgatatttccctgcgcgttggttggttactactggtggccagg 372
Db 343 --GAAAGGGTCACAGTGGCTGA-----CCCCCTCGTTCTGCTTTGACTACTGGGGCCACG 395
QY 373 ggaccagggtcaccgtctcctcaggatc 400
Db 396 GTACCCCTGGTCACCCGTCCTCCTCAGCCTC 423

```



QY	73	cctgtgcagctcgggacgcgcaccagtggtcatggtcactatggtatgggctggttcc	132
Db	143	CTGTGTCAGCTCTGGATTACCTTTAGCAG-----CTATGCCATGAGCTGGGTCC	193
QY	133	gccaggttccagggaaggagcgtgagttgtgcagctattaggtggagtggttaaagaga	192
Db	194	GCCAGGCTCCAGGAAGGGCTGGAGTGGTCTCAGCTATTAGTGGTAGGTAGCA	253
QY	193	catggtataaagactccgtgaaggccgattcaccatctccagagagataaacgccaagacta	252
Db	254	CATACTACGACACTCCGTGAAGGCCGGTTCACCATCTCCAGAGACAAATCCAAGACA	313
QY	253	cggtttatctgcaaatgaacagcctgaacctgaagatacgcgcgtttattattgtgccg	312
Db	314	CGCTGTATCTGCAAAATGAACAGCCTGAGAGCCGAGGACACGCGCGTATATTACTGTGCGA	373
QY	313	ctcgaccggt 322	
Db	374	CCCACCCGGT 383	
RESULT	11		
AW401799			
LOCUS	AW401799	559 bp	linear
DEFINITION	UI-HF-BK0-aag-g-06-0-UI.r1 NIH_MGC_36	Homo sapiens	CDNA clone
ACCESSION	AW401799		
VERSION	AW401799.1	GI:6920485	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	1 (bases 1 to 559)		
TITLE	NIH-MGC http://mgc.nci.nih.gov/.		
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)		
COMMENT	Unpublished (1999)		
	Contact: Robert Strausberg, Ph.D.		
	Email: cgapbs-r@mail.nih.gov		
	Eco RI site shown at the beginning of the sequence.		
	Tissue Procurement: Louis M. Staudt, M.D., Ph.D.		
	CDNA Library Preparation: M.B. Soares Lab		
	CDNA Library Arrayed by: M.B. Soares Lab		
	DNA Sequencing by: M.B. Soares Lab		
	Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/bbrp/image/image.html		
	Seq primer: M13 Forward.		
FEATURES	Location/Qualifiers		
source	1..559		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone="IMAGE:3053770"		
	/clone_lib="NIH_MGC_36"		
	/tissue_type="lymph"		
	/cell_type="germinal center B cells"		
	/cell_line="MGC85"		
	/lab_host="DH10B (LTI)"		
	/note="Vector: pT7T3-Pac; Site_1: NotI; Site_2: Eco RI; Constructed from size fractionated cytoplasmic mRNA (0.5-1.5kb). Directionally cloned. Cells provided by Louis M. Staudt, Ph.D. Library preparation by Maria de Fatima Bonaldo, Ph.D. and M. Bento Soares, Ph.D."		
BASE COUNT	116 a 150 c 158 g 135 t		
ORIGIN			
Query Match	32.7%	Score 182.4;	DB 9; Length 559;
Best Local Similarity	69.1%	Pred. NO. 1.7e-43;	
Matches 268; Conservative	0;	Mismatches 111;	Indels 9; Gaps 1;
QY	13	aggtgcagctgcaggagggaggattggtgcaggtggggctctctgagactct	72

Db	128	AGGTGCAGCTGCTGGCGCTCTGGGGACGCGCTGGTACGCGCTGGGGGCTCCCTGAGACTCT	187
QY	73	cctgtgcagctcgggacgcgcaccagtggtcatggtcactatggtatgggctggttcc	132
Db	188	CCTGTGTAGTCTCTGGATTACCC-----TTTGATGATTATGGCATGAGCTGGGTCC	238
QY	133	gccaggttccagggaaggagcgtgagttgtgcagctattaggtggagtggttaaagaga	192
Db	239	GCCAACTTCCAGGAAGGGCTGGAGTGGGTCTCTGGCATTAATFGGAATGGTGGCACTA	298
QY	193	catggtataaagactccgtgaaggccgattcaccatctccagagagataaacgccaagacta	252
Db	299	CAGCTTATGCAGACTCTGTGAAAGGCCGATTACCATCTCCAGAGACACGCCAAGAACT	358
QY	253	cggtttatctgcaaatgaacagcctgaacctgaagatacgcgcgtttattattgtgccg	312
Db	359	CCCTATTTCTGCAAAATGAACAGTCTGAAAGGCCGAGGACACGGCCTTCTATTACTGTGCGA	418
QY	313	ctcgaccggtcgcggtggatgatatttccctgcgcgttgggtttgactactggggccagg	372
Db	419	GAGTGCACCAAGGCTATGACAGTGCCAACTCTTACTACTTTGACTACTGGGGCCAGG	478
QY	373	ggaccaggtcacgcgtctcctcaggatc	400
Db	479	GAACCCCTGGTCACCGTCTCCTCAGCCTC	506
RESULT	12		
BG745331			
LOCUS	BG745331	759 bp	linear
DEFINITION	602723655F1 NIH_MGC_113	Homo sapiens	CDNA clone
ACCESSION	BG745331		
VERSION	BG745331.1	GI:14055984	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	1 (bases 1 to 759)		
TITLE	NIH-MGC http://mgc.nci.nih.gov/.		
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)		
COMMENT	Unpublished (1999)		
	Contact: Robert Strausberg, Ph.D.		
	Email: cgapbs-r@mail.nih.gov		
	Tissue Procurement: Dr. Mark Watson		
	CDNA Library Preparation: Ling Hong/Rubin Laboratory		
	CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)		
	DNA Sequencing by: Incyte Genomics, Inc.		
	Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov		
	Plate: LLCM1690	row: n	column: 14
	High quality sequence stop: 759.		
FEATURES	Location/Qualifiers		
source	1..759		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone="IMAGE:4850125"		
	/clone_lib="NIH_MGC_113"		
	/lab_host="DH10B (phage-resistant)"		
	/note="Organ: spleen; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."		
BASE COUNT	158 a 215 c 215 g 171 t		
ORIGIN			





```

/tissue_type="lymph"
/cell_type="germinal center B cells"
/cell_line="MGC85"
/lab_host="DH10B (LTI)"
/Note="Vector: pT7T3-Pac; Site_1: NotI; Site_2: Eco RI;
Constructed from size fractionated cytoplasmic mRNA
(0.5-1.5kb). Directionally cloned. Cells provided by Louis
M. Staudt, Ph.D. Library preparation by Maria de Fatima
Bonaldo, Ph.D. and M. Bento Soares, Ph.D."
BASE COUNT      98 a  127 c  143 g  103 t
ORIGIN
Query Match      32.6%; Score 182; DB 9; Length 471;
Best Local Similarity 71.4%; Pred. No. 2e-43;
Matches 277; Conservative 0; Mismatches 90; Indels 21; Gaps 2;

QY 13 aggtgcagctgcaggagtcaggggaggattggtgcaggctggggctctctgcagactct 72
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 77 AGTGCAGCTGGTGGAGTCTGGGGGAGGCCTGGTCAAGCCTGGGGGTCCCTGCAGACTCT 136

QY 73 cctgtgcagcctcgggacgcgccaccagtggtcatggtcactatggtatggcgctgtcc 132
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 137 CCTGTGCAGCCTCTGGATTACCTTCAGTAG-----CGATAGCATGAACCTGGGTCC 187

QY 133 gccaggttccagggaagcagcgtgagttgtgcagctattagtgagggtgtaaagaga 192
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 188 GCCAGGCTCCAGGGAAGGGCTGGAGTGGGTCTCATCCATTAGTAGTAGTGGTTTCA 247

QY 193 catggtataaagactccgtgaaggccgattccaccatccagagataacgccaaagacta 252
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 248 TATACTACGCAGACTCAGTGAAGGGCCGATTACCATCTCCAGAGACAACGCCAAGACT 307

QY 253 cggtttatctgcaaatgaacagcctgaaacctgaagatacggccgtttattattgtgcg 312
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 308 CACTGTATCTGCAATGAACAGCCTGAGAGCCGAGGACACGGCTGTGTATTACTGTGC-- 365

QY 313 ctgcaccggtccgctggtgatgatattccctgcgcggttggtggttgcactactggggccag 372
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 366 -----GAGAGCCCATAGTAGCAGCTGGCCACGCACGACTGACTACTGGGGCCAGG 415

QY 373 ggaccaggtcaccgtctcctcaggatc 400
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 416 GAACCTGGTCACCGCTCCTCAGCCTC 443

RESULT 15
AW401728
LOCUS      AW401728          516 bp      mRNA      linear      EST 16-FEB-2000
DEFINITION      UI-HF-BK0-aaf-f-12-0-UI.r1 NIH_MGC_36 Homo sapiens cDNA clone
IMAGE:3053711 5', mRNA sequence.
ACCESSION      AW401728
VERSION        AW401728.1 GI:6920414
KEYWORDS        EST.
SOURCE          human.
ORGANISM        Homo sapiens
                Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
                1 (bases 1 to 516)
                NIH-MGC http://mgc.nci.nih.gov/.
                National Institutes of Health, Mammalian Gene Collection (MGC)
                Unpublished (1999)
                Contact: Robert Strausberg, Ph.D.
                Email: cgapbs-r@mail.nih.gov
                Eco RI site shown at the beginning of the sequence.
                Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
                CDNA Library Preparation: M.B. Soares Lab
                CDNA Library Arrayed by: M.B. Soares Lab
                DNA Sequencing by: M.B. Soares Lab
                Clone distribution: MGC clone distribution information can be
                found through the I.M.A.G.E. Consortium/LLNL at:
                www-bio.llnl.gov/bbrp/image/image.html
                Seq primer: M13 Forward.

```

OM of: US-09-742-690-1 to: SPTREMBL\_19:\* out\_format : pfs  
Date: Aug 13, 2002 1:53 PM  
About: Results were produced by the GenCore software, version 4.5,  
Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:  
-MODEL=frame-n2p.model -DEV=xlp  
-Q=/cgn2\_1/USPTO\_spool/US09742690/runat\_12082002\_094915\_166/app\_query.fasta\_1.623  
-DB=SPTREMBL\_19 -QFMT=fastan -SUFFIX=rspt -GAPOP=12.000  
-GAPEXT=4.000 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000  
-QGAPOP=4.500 -QGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500  
-FGAPOP=6.000 -FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500  
-DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=blossum62  
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR\_SCORE=pct  
-THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs  
-NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09742690@cgn1\_1\_132 -NCPU=6 -ICPU=3 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -NO\_XLPXY -WAIT -THREADS=1

Search information block:

Query: US-09-742-690-1  
Query length: 558  
Database: SPTREMBL\_19:\*  
Database sequences: 562222  
Database length: 172994929  
Search time (sec): 101.840000

score_list:	Sequence	Strd	Orig	zScore	EScore	Len	Documentation
	sp_human:Q96BB9	+	402.50	698.17	2.6e-31	597	! Q96bb9 homo sapiens (human).
	sp_human:Q9UL90	+	389.50	690.98	3.4e-30	113	! Q9ul90 homo sapiens (human).
	sp_human:Q920E7	+	379.50	672.65	3.4e-29	119	! Q920e7 mus musculus (mouse).
	sp_human:Q9HCC1	+	376.00	666.99	7.5e-29	112	! Q9hcc1 homo sapiens (human).
	sp_human:Q9UL93	+	369.50	655.06	3.3e-28	116	! Q9ul93 homo sapiens (human).
	sp_human:Q9UL71	+	369.50	654.65	3.4e-28	121	! Q9ul71 homo sapiens (human).
	sp_human:Q99KA4	+	369.00	640.38	5.2e-28	487	! Q99ka4 mus musculus (mouse).
	sp_human:Q9Y509	+	368.00	650.11	4.9e-28	147	! Q9y509 homo sapiens (human).
	sp_human:Q9UL72	+	365.00	646.87	9.3e-28	118	! Q9ul72 homo sapiens (human).
	sp_human:Q96K68	+	364.00	631.32	1.6e-27	494	! Q96k68 homo sapiens (human).
	sp_human:Q9UL91	+	362.00	641.52	1.9e-27	118	! Q9ul91 homo sapiens (human).
	sp_human:Q9UL84	+	360.00	637.63	3.0e-27	122	! Q9ul84 homo sapiens (human).
	sp_human:Q91WP5	+	353.50	612.89	1.8e-26	479	! Q91wp5 mus musculus (mouse).
	sp_human:Q91Z07	+	346.50	600.27	8.9e-26	486	! Q91z07 mus musculus (mouse).
	sp_human:Q91Z05	+	342.00	592.50	2.5e-25	473	! Q91z05 mus musculus (mouse).
	sp_human:Q9UL86	+	338.50	601.69	3.8e-25	95	! Q9ul86 homo sapiens (human).
	sp_human:Q9UL88	+	327.50	578.98	5.1e-24	131	! Q9ul88 homo sapiens (human).
	sp_human:Q9QYF0	+	327.50	571.08	6.2e-24	298	! Q9qyf0 mus musculus (mouse).
	sp_human:Q91XE1	+	325.50	562.93	1.1e-23	480	! Q91xe1 mus musculus (mouse).
	sp_mammal:Q9N1A4	+	305.00	527.27	1.2e-21	437	! Q9n1a4 mus musculus (mouse).
	sp_mammal:Q9N0W6	+	295.00	521.54	8.5e-21	124	! Q9n0w6 oryctolagus cuniculus (mouse).
	sp_mammal:Q9N0W4	+	293.00	517.98	1.3e-20	124	! Q9n0w4 oryctolagus cuniculus (mouse).
	sp_human:Q9UGP3	+	289.50	512.71	2.9e-20	112	! Q9ugp3 homo sapiens (human).
	sp_human:Q99NG4	+	285.00	503.94	8.3e-20	121	! Q99ng4 mus musculus (mouse).
	sp_human:Q91X92	+	280.00	481.74	3.6e-19	482	! Q91x92 mus musculus (mouse).
	sp_human:Q9UL87	+	279.50	495.59	2.8e-19	104	! Q9ul87 homo sapiens (human).
	sp_human:Q921A6	+	279.00	486.62	3.9e-19	241	! Q921a6 mus musculus (mouse).
	sp_human:Q9UL92	+	276.00	487.66	6.6e-19	124	! Q9ul92 homo sapiens (human).
	sp_human:Q9UL75	+	275.00	486.03	8.2e-19	122	! Q9ul75 homo sapiens (human).
	sp_human:Q96KX8	+	268.00	460.06	5.6e-18	496	! Q96kx8 homo sapiens (human).
	sp_human:Q95973	+	267.00	469.77	5.4e-18	150	! Q95973 homo sapiens (human).
	sp_human:Q96EY0	+	264.00	450.89	1.5e-17	613	! Q96ey0 homo sapiens (human).
	sp_human:Q924Q2	+	262.50	462.28	1.5e-17	142	! Q924q2 mus musculus (mouse).
	sp_human:Q991C4	+	261.00	448.24	2.8e-17	463	! Q991c4 mus musculus (mouse).
	sp_human:Q924R4	+	259.00	455.83	3.3e-17	145	! Q924r4 mus musculus (mouse).
	sp_human:Q924R1	+	258.00	454.05	4.2e-17	145	! Q924r1 mus musculus (mouse).
	sp_human:Q924P7	+	258.00	454.05	4.2e-17	145	! Q924p7 mus musculus (mouse).
	sp_human:Q96QSO	+	258.00	453.16	4.3e-17	159	! Q96qso homo sapiens (human).
	sp_human:Q924Q1	+	257.50	453.36	4.7e-17	142	! Q924q1 mus musculus (mouse).
	sp_human:Q921C4	+	257.00	454.25	5.0e-17	118	! Q921c4 mus musculus (mouse).

sp_rodent:Q924Q8	+	256.50	451.31	5.9e-17	146	! Q924q8 mus musculus (mouse).
sp_rodent:Q924Q3	+	256.50	451.31	5.9e-17	146	! Q924q3 mus musculus (mouse).
sp_rodent:Q9QXF0	+	255.50	451.65	7.0e-17	117	! Q9qxf0 mus musculus (mouse).
sp_human:Q9UL73	+	255.50	451.49	7.1e-17	119	! Q9ul73 homo sapiens (human).
sp_rodent:Q99LA6	+	255.50	438.01	9.8e-17	484	! Q99la6 mus musculus (mouse).
seq_name: sp_human:Q96BB9						
seq_documentation_block:						
ID	Q96BB9	PRELIMINARY;	PRT;	597	AA.	
AC	Q96BB9;					
DT	01-DEC-2001	(TREMBLrel. 19, Created)				
DT	01-DEC-2001	(TREMBLrel. 19, Last sequence update)				
DT	01-DEC-2001	(TREMBLrel. 19, Last annotation update)				
DE	HYPOTHETICAL 65.0 KDA PROTEIN.					
OS	Homo sapiens (Human).					
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.					
OX	NCBI_TaxID=9606;					
RN	[1]					
RP	SEQUENCE FROM N.A.					
RC	TISSUE=PRIMARY B-CELLS FROM TONSILS;					
RA	Strausberg R.;					
RL	Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.					
DR	EMBL; BC015760; AAH15760.1;					
KW	Hypothetical protein.					
SQ	SEQUENCE 597 AA; 65039 MW; 4FCA3AD8ECE263D9 CRC64;					

alignment\_scores:

Quality: 402.50 Length: 130  
Ratio: 3.946 Gaps: 1  
Percent Similarity: 78.462 Percent Identity: 65.385

alignment\_block:

US-09-742-690-1 x Q96BB9 ..

Align seg 1/1 to: Q96BB9 from: 1 to: 597

12	CAGGTGCAGCTGAGGAGTCCAGGGGAGGATTGGTCAGGCTGGGGGCTC	61
:		
20	GluValGlnLeuLeuGluSerGlyGlyGlyLeuValGlnProGlyGlyse	36
62	TCTGAGACTCTCTGTGTCAGCTCGGACGGCCAGCAGTGGTTCATGGTC	111
:		
36	rLeuArgLeuSerCysAlaAlaSerGlyPheSerPheSerSer.....	50
112	ACTATGGTATGGGTGGTTCGCCAGGTTCCAGGGAAGGAGCGTGAGTTT	161
:		
51	..TyrAlaMetAsnTrpValArgGlnAlaProGlyLysGlyLeuGluTrp	66
162	GTGCGCAGCTATTAGTGGAGTGGTAAAGAGACATGGTATAAGACTCCGT	211
:		
67	ValSerAlaIleSerGlySerGlyGlySerThrTyrTyrAlaAspSera	83
212	GAAGGGCCGATTACCATCTCCAGAGATAACGCCAAGACTACGGTTTATC	261
:		
83	LLysGlyArgPheThrIleSerArgAspAsnSerArgAspThrLeuTyrL	100
262	TGCAAAATGAACAGCCTGAAACCTGAAGATACGGCGTTTATTATTGTGCC	311
:		
100	euGlnMetAsnSerLeuArgAlaGluAspThrAlaValTyrTyrCysAla	116
312	GCTCGACCGTCCCGGTGGATGATATTTCCCTGCCGTTGGTTTGACTA	361
:		
117	LysAspProArgGlyTyrSerAlaSerGlyAsnTyrThrArgGluAspTy	133
362	CTGGGGCCAGGGACCCAGGTACCGTCTCTCAGGATCT	401
:		
133	rTrpGlyGlnGlyThrLeuValThrValSerSerGlySer	146

seq\_name: sp\_human:Q9UL90

```
seq_documentation_block:
ID Q9UL90 PRELIMINARY; PRT; 113 AA.
AC Q9UL90;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE MYOSIN-REACTIVE IMMUNOGLOBULIN HEAVY CHAIN VARIABLE REGION
DE (FRAGMENT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98277139; PubMed=9614934;
RA Wu X., Liu B., Van der Merwe P.L., Kalis N.N., Berney S.M.,
RA Young D.C.;
RT "Myosin-reactive autoantibodies in rheumatic carditis and normal
RT fetus.";
RL Clin. Immunol. Immunopathol. 87:184-192(1998).
DR EMBL; AF035024; AAD56260.1; -.
DR HSSP; P01772; 2FB4.
DR InterPro; IPR003006; Ig_MHC.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; Igv; 1.
FT NON_TER 1
FT NON_TER 113
SQ SEQUENCE 113 AA; 12437 MW; ED57FDD19086D07F CRC64;
```

```
alignment_scores:
Quality: 389.50 Length: 128
Ratio: 3.974 Gaps: 2
Percent Similarity: 76.562 Percent Identity: 63.281
```

```
alignment_block:
US-09-742-690-1 x Q9UL90 ..
```

Align seg 1/1 to: Q9UL90 from: 1 to: 113

```
12 CAGGTGCAGCTGCAGGAGTCAGGGGGAGGATTGGTCAGGCTGGGGGCTC 61
:::||||| ||||| ||||| ||||| ||||| ||||| |||||
1 GluValGlnLeuValGluSerGlyGlyValValGlnProGlyGlySe 17

62 TCTGAGACTCTCCTGTGAGCCCTCGGACGCGCCACAGTGGTCATGGTC 111
||||| ||||| ||||| ||||| ||||| ||||| |||||
17 rLeuArgLeuSerCysAlaAlaSerGlyPheThrPheSerSer..... 31

112 ACTATGGTATGGCTGTTCCGCCAGGTTCCAGGGAAGAGCGTGAGTTT 161
||||| ||||| ||||| ||||| ||||| ||||| |||||
32 ..TyrGlyMetHisTrpValArgGlnAlaProGlyLysGlyLeuGluTrp 47

162 GTCGCAGCTATTAGGTGGAGTGGTAAAGAGACATGGTATAAGACTCCGT 211
||||| ||||| ||||| ||||| ||||| ||||| |||||
48 ValAlaPheIleArgTyrAspGlySerAsnLysTyrTyrAlaAspSerVa 64

212 GAAGGGCCGATTACCATCTCCAGAGATAACGCCAAGACTACGGTTTATC 261
||||| ||||| ||||| ||||| ||||| ||||| |||||
64 llysGlyArgPheThrIleSerArgAspAsnSerLysAsnThrLeuTyrL 81

262 TGCAAATGAACAGCCTGAACCTGAAGATACGGCCGTTTATTATTGTGCC 311
||||| ||||| ||||| ||||| ||||| ||||| |||||
81 euGlnMetAsnSerLeuArgAlaGluAspThrAlaValTyrTyrCysAla 97

312 GCTCGACCGGTCGCGGTGGATGATATTTCCCTGCCGTTGGGTTTGACTA 361
:::|::|
98 Lys.....AspLeuAsnTy 102
```

```
362 CTGGGGCCAGGGACCCAGGTACCGTCTCCTCA 395
||||| ||||| ||||| ||||| ||||| ||||| |||||
102 rTrpGlyGlnGlyThrLeuValThrValSerSer 113
```

```
seq_name: sp_rodent:Q920E7

seq_documentation_block:
ID Q920E7 PRELIMINARY; PRT; 119 AA.
AC Q920E7;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE PTERIN-MIMICKING ANTI-IDIOPOPE HEAVY CHAIN VARIABLE REGION
DE (FRAGMENT).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Atkin J.D., Iape A., Jennings I.G., Horaitis O., Cotton R.G.H.;
RT "Definition of the Idiotope of Pterin-Mimicking Antibodies Expressed
RT in Mammalian Cells.";
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF307937; AAL09421.1; -.
FT NON_TER 1
FT NON_TER 119
SQ SEQUENCE 119 AA; 13025 MW; F6E904044381CA7C CRC64;
```

```
alignment_scores:
Quality: 379.50 Length: 128
Ratio: 3.757 Gaps: 3
Percent Similarity: 78.906 Percent Identity: 64.062
```

```
alignment_block:
US-09-742-690-1 x Q920E7 ..
```

Align seg 1/1 to: Q920E7 from: 1 to: 119

```
12 CAGGTGCAGCTGCAGGAGTCAGGGGGAGGATTGGTCAGGCTGGGGGCTC 61
:::||||| ||||| ||||| ||||| ||||| ||||| |||||
1 GluValGlnLeuValGluSerGlyGlyAspLeuValLysProGlyGlySe 17

62 TCTGAGACTCTCCTGTGAGCCCTCGGACGCGCCACAGTGGTCATGGTC 111
||||| ||||| ||||| ||||| ||||| ||||| |||||
17 rLeuLysLeuSerCysAlaAlaSerGlyPheThrPheSerSer..... 31

112 ACTATGGTATGGCTGTTCCGCCAGGTTCCAGGGAAGAGCGTGAGTTT 161
||||| ||||| ||||| ||||| ||||| ||||| |||||
32 ..TyrGlyMetSerTrpValArgGlnThrProAspLysArgLeuGluTrp 47

162 GTCGCAGCTATTAGGTGGAGTGGTAAAGAGACATGGTATAAGACTCCGT 211
||||| ||||| ||||| ||||| ||||| ||||| |||||
48 ValAlaThrIleSerSerGlyGlySerTyrThrTyrTyrProAspSerVa 64

212 GAAGGGCCGATTACCATCTCCAGAGATAACGCCAAGACTACGGTTTATC 261
||||| ||||| ||||| ||||| ||||| ||||| |||||
64 llysGlyArgPheThrIleSerArgAspAsnAlaLysAsnThrLeuTyrL 81

262 TGCAAATGAACAGCCTGAACCTGAAGATACGGCCGTTTATTATTGTGCC 311
||||| ||||| ||||| ||||| ||||| ||||| |||||
81 euGlnMetSerSerLeuLysSerGluAspThrAlaMetTyrTyrCysAla 97

312 GCTCGACCGGTCGCGGTGGATGATATTTCCCTGCCGTTGGGTTTGACTA 361
||| ||| ||| ::| ||||| |||
98 .....ArgHisGlyAspTyrAsp.....ValGlyPheAlaTy 108
```

```
362 CTGGGGCCAGGGACCCAGGTACCGTCTCCTCA 395
||||| ||||| ||||| ||||| ||||| ||||| |||||
108 rTrpGlyGlnGlyThrLeuValThrValSerAla 119
```

seq\_name: sp\_human:Q9HCC1

```
seq_documentation_block:
ID Q9HCC1 PRELIMINARY; PRT; 112 AA.
AC Q9HCC1;
```





```
seq_name: sp_human:Q9UL71
seq_documentation_block:
ID Q9UL71 PRELIMINARY; PRT; 121 AA.
AC Q9UL71;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE MYOSIN-REACTIVE IMMUNOGLOBULIN HEAVY CHAIN VARIABLE REGION
DE (FRAGMENT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98277139; PubMed=9614934;
RA Wu X., Liu B., Van der Merwe P.L., Kalis N.N., Berney S.M.,
RA Young D.C.;
RT "Myosin-reactive autoantibodies in rheumatic carditis and normal
RT fetus.";
RL Clin. Immunol. Immunopathol. 87:184-192(1998).
DR EMBL; AF035043; AAD56279.1; -.
DR HSSP; P01772; 2FB4.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; ig; 1.
DR SMART; SM00406; IGV; 1.
FT NON_TER 1
FT NON_TER 121
FT SEQUENCE 121 AA; 13154 MW; 2F045CCFA5D50736 CRC64;

alignment_scores:
Quality: 369.50 Length: 128
Ratio: 3.770 Gaps: 2
Percent Similarity: 76.562 Percent Identity: 60.938

alignment_block:
US-09-742-690-1 x Q9UL71 ..
Align seg 1/1 to: Q9UL71 from: 1 to: 121

12 CAGGTGCAGCTGCAGGAGTGCAGGGGAGGAGTGGTGCAGGCTGGGGGCTC 61
:::||||| ||||| ||||| ||||| ||||| ||||| |||||
1 GluValGlnLeuValGluSerGlyGlyValValGlnProGlyGlySe 17

62 TCTGAGACTCTCCTGTGTCAGCCTCGGACGCGCCACCAGTGGTGTATGTC 111
||||| ||||| ||||| ||||| ||||| ||||| |||||
17 rLeuArgLeuPheCysAlaAlaSerGlyPheThrPheAspGly..... 31

112 ACTATGGTATGGCTGGTTCGCCAGGTTCCAGGGAAGAGCGTGAGTTT 161
||||| ||||| ||||| ||||| ||||| ||||| |||||
32 ..TyrAlaMetHisTrpValArgGlnAlaProGlyLysGlyLeuGluTrp 47

162 GTCGCAGCTATAGGTGGAGTGGTAAAGAGACATGGTATAAGACTCCGT 211
||||| ||||| ||||| ||||| ||||| ||||| |||||
48 ValSerLeuIleSerGlyAspGlySerThrThrTyrAlaAspSerVa 64

212 GAAGGGCGGATTCACCATCTCCAGAGATAACGCCAAGACTACGGTTATC 261
||||| ||||| ||||| ||||| ||||| ||||| |||||
64 lLysGlyArgPheThrIleSerArgAspAsnSerLysAsnSerLeuTyrL 81

262 TGCAATGAACAGCCTGAAACCTGAAGATACGGCGGTTTATTATTGTGCC 311
||||| ||||| ||||| ||||| ||||| ||||| |||||
81 euGlnMetAsnSerLeuArgAlaGluAspThrAlaLeuTyrTyrCysAla 97

312 GCTCGACCGGTCGGGTGGATGATATTTCCCTGCCGGTTGGTTGACTA 361
|||||
98 LysGlyLysVal.....ThrThrIleTyrAspArgPheAspI1 110

362 CTGGGGCAGGGGACCCAGGTACCGTCTCCTCA 395
||||| ||||| ||||| ||||| ||||| ||||| |||||
```

```
110 eTrpGlyGlnGlyThrMetValThrValSerSer 121
seq_name: sp_rodent:Q99KA4
seq_documentation_block:
ID Q99KA4 PRELIMINARY; PRT; 487 AA.
AC Q99KA4;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE HYPOTHETICAL 52.6 KDA PROTEIN.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC004786; AAH04786.1; -.
DR HSSP; P01810; 2FBJ.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR003597; Ig_cl.
DR InterPro; IPR003600; Ig_like.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; ig; 4.
DR SMART; SM00409; IG; 3.
DR SMART; SM00407; IGcl; 3.
DR SMART; SM00406; IGV; 1.
DR SMART; SM00410; IG_like; 1.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_2.
KW Hypothetical protein.
SQ SEQUENCE 487 AA; 52554 MW; 7DC8E96DB333077B CRC64;

alignment_scores:
Quality: 369.00 Length: 134
Ratio: 3.618 Gaps: 3
Percent Similarity: 76.119 Percent Identity: 59.701

alignment_block:
US-09-742-690-1 x Q99KA4 ..
Align seg 1/1 to: Q99KA4 from: 1 to: 487

12 CAGGTGCAGCTGCAGGAGTGCAGGGGAGGAGTGGTGCAGGCTGGGGGCTC 61
:::||||| ||||| ||||| ||||| ||||| ||||| |||||
20 GluValGlnLeuValGluSerGlyGlyGlyLeuValLysProGlyGlySe 36

62 TCTGAGACTCTCCTGTGTCAGCCTCGGACGCGCCACCAGTGGTGTATGTC 111
||||| ||||| ||||| ||||| ||||| ||||| |||||
36 rLeuLysLeuSerCysAlaAlaSerGlyPheThrPheSerSer..... 50

112 ACTATGGTATGGCTGGTTCGCCAGGTTCCAGGGAAGAGCGTGAGTTT 161
||||| ||||| ||||| ||||| ||||| ||||| |||||
51 ..TyrAlaMetSerTrpValArgGlnThrProGluLysArgLeuGluTrp 66

162 GTCGCAGCTATAGGTGGAGTGGTAAAGAGACATGGTATAAGACTCCGT 211
||||| ||||| ||||| ||||| ||||| ||||| |||||
67 ValAlaThrIleSerAspGlyGlySerThrThrTyrProAspAsnVa 83

212 GAAGGGCGGATTCACCATCTCCAGAGATAACGCCAAGACTACGGTTATC 261
||||| ||||| ||||| ||||| ||||| ||||| |||||
83 lLysGlyArgPheThrIleSerArgAspAsnAlaLysAsnLeuTyrL 100

262 TGCAATGAACAGCCTGAAACCTGAAGATACGGCGGTTTATTATTGTGCC 311
||||| ||||| ||||| ||||| ||||| ||||| |||||
100 euGlnMetSerHisLeuLysSerGluAspThrAlaMetTyrTyrCysAla 116

312 GCTCGACCGGTCGGGTGGATGATATTTCCCTGCCGGTTGGG..... 353
|||||
117 Arg.....AspMetGlyGlySerProTyrGlyGlyTyrSe 128
```

354 ....TTTGACTACTGGGGCCAGGGGACCCAGGTCACCGTCTCTCCTCAGGAT 399  
|||||  
128 rArgpheaSpTyrTrpGlyGlnGlyThrThrIleThrValSerSerGluS 145  
400 CT 401  
||  
145 er 145

seq\_name: sp\_human:Q9Y509

seq\_documentation\_block:  
ID Q9Y509 PRELIMINARY; PRT; 147 AA.  
AC Q9Y509;  
DT 01-NOV-1999 (TReMBLrel. 12, Created)  
DT 01-NOV-1999 (TReMBLrel. 12, Last sequence update)  
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)  
DE VH3 PROTEIN (FRAGMENT).  
GN VH3.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=96071149; PubMed=7475288;  
RA Cao J., Vescio R.A., Rettig M.B., Hong C.H., Kim A., Lee J.C.,  
RA Lichtenstein A.K., Berenson J.R.;  
RT "A CD10-positive subset of malignant cells is identified in multiple  
myeloma using PCR with patient-specific immunoglobulin gene primers.";  
RL Leukemia 9:1948-1953(1995).  
DR EMBL; S80860; AAD14339.1; -.  
DR HSSP; P01772; 2FB4.  
DR InterPro; IPR003006; Ig\_MHC.  
DR InterPro; IPR003596; Ig\_v.  
DR Pfam; PF00047; ig; 1.  
DR SMART; SM00406; Igv; 1.  
FT NON\_TER 1  
SQ SEQUENCE 147 AA; 15768 MW; 8489FCAAA7BC925C CRC64;

alignment\_scores:  
Quality: 368.00 Length: 131  
Ratio: 3.644 Gaps: 2  
Percent Similarity: 77.099 Percent Identity: 60.305  
alignment\_block:  
US-09-742-690-1 x Q9Y509 ..  
Align seg 1/1 to: Q9Y509 from: 1 to: 147

12 CAGGTGCAGCTGCAGGAGTGCAGGGGAGGATTGGTCAGGCTGGGGGCTC 61  
|||||  
1 GlnValHisLeuValGluSerGlyGlyValValGlnProGlyLysSe 17  
62 TCTGAGACTCTCTGTGCAGCCTCGGACGCGCCACCAAGTGGTCATGGTC 111  
|||||  
17 rLeuArgLeuSerCysGluAlaSerGlyPheThrPheSer.....T 31  
112 ACTATGGTATGGCTGGTTCCGCCAGGTTCCAGGGAAGGAGCGTGAGTTT 161  
|||||  
31 hrTyrGlyMetSerTirpValArgGlnAlaProGlyLysGlyLeuAspTrp 47  
162 GTCGCAGCTATTAGTGGAGTGGTAAAGACACATGGTATAAAGACTCCGT 211  
|||||  
48 ValAlaLeuIleSerTyrAspGlySerThrGlnTyrTyrAlaGlySerVa 64  
212 GAAGGGCCGATTACCATCTCCAGAGATAACGCCAAGACTACGGTTTATC 261  
|||||  
64 lLysGlyArgPheThrIleSerArgAspAsnSerLysAsnThrLeuTyrL 81  
262 TGCAAATGAACAGCCTGAAACCTGAAGATACGGCCGTTTATTATTGTGCC 311  
|||||

81 euGlnMetThrSerLeuArgValGluAspThrAlaValTyrTyrCysAla 97  
312 GCTCGACCGGTCCGGTGGATGATATTCC...CTGCCGTTGGGTTGA 358  
|||||  
98 LysAspGlyAsnTyrPheAspSerValGlyTyrTyrTyrAlaGlyIleAs 114  
359 CTACTGGGCGCAGGGACCCAGGTCACCGTCTCTCCTCAGGATCT 401  
|||||  
114 pTyrTrpGlyGlnGlyThrLeuValThrValSerSerAlaSer 128  
seq\_name: sp\_human:Q9UL72  
seq\_documentation\_block:  
ID Q9UL72 PRELIMINARY; PRT; 118 AA.  
AC Q9UL72;  
DT 01-MAY-2000 (TReMBLrel. 13, Created)  
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)  
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)  
DE MYOSIN-REACTIVE IMMUNOGLOBULIN HEAVY CHAIN VARIABLE REGION  
(FRAGMENT).  
DE Homo sapiens (Human).  
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=98277139; PubMed=9614934;  
RA Wu X., Liu B., Van der Merwe P.L., Kalis N.N., Berney S.M.,  
RA Young D.C.;  
RT "Myosin-reactive autoantibodies in rheumatic carditis and normal  
fetuses";  
RL Clin. Immunol. Immunopathol. 87:184-192(1998).  
DR EMBL; AF035042; AAD56278.1; -.  
DR HSSP; P01772; 2FB4.  
DR InterPro; IPR003006; Ig\_MHC.  
DR InterPro; IPR003596; Ig\_v.  
DR Pfam; PF00047; ig; 1.  
DR SMART; SM00406; Igv; 1.  
FT NON\_TER 1  
FT NON\_TER 118  
SQ SEQUENCE 118 AA; 12872 MW; B4D1A5944B2D5CCA CRC64;

alignment\_scores:  
Quality: 365.00 Length: 128  
Ratio: 3.763 Gaps: 3  
Percent Similarity: 75.781 Percent Identity: 62.500  
alignment\_block:  
US-09-742-690-1 x Q9UL72 ..  
Align seg 1/1 to: Q9UL72 from: 1 to: 118

12 CAGGTGCAGCTGCAGGAGTGCAGGGGAGGATTGGTCAGGCTGGGGGCTC 61  
:::|||||  
1 GluValGlnLeuValGluSerGlyGlyGlyLeuValGlnProGlyGlySe 17  
62 TCTGAGACTCTCTGTGCAGCCTCGGACGCGCCACCAAGTGGTCATGGTC 111  
|||||  
17 rLeuArgLeuSerCysAlaAlaSerGlyPheThrValSerSerAsn.... 32  
112 ACTATGGTATGGCTGGTTCCGCCAGGTTCCAGGGAAGGAGCGTGAGTTT 161  
|||||  
33 .....TyrMetAsnTrpValArgGlnAlaProGlyLysGlyLeuGluSer 47  
162 GTCGCAGCTATTAGTGGAGTGGTAAAGACACATGGTATAAAGACTCCGT 211  
|||  
48 Val...SerValThrTyrSerGlySerSerTyrTyrAlaAspSerVa 63  
212 GAAGGGCCGATTACCATCTCCAGAGATAACGCCAAGACTACGGTTTATC 261  
|||||  
63 lLysGlyArgPheThrIleSerArgAspAsnSerLysAsnThrLeuTyrL 80

```
262 TGCAAAATGAACAGCCTGAAACCTGAAGATACGGCGTTTATTATTGTGCC 311
|||||
80 euGlnMetAsnSerLeuArgAlaGluAspThrAlaPheTyrTyrCysAla 96
|||||
312 GCTCGACCGTCCGCGTGGATGATATTTCCCTGCGCGTTGGGTTTGACTA 361
|||
97 Arg.....AspArgPheGlyGluPheLeuPheAspTyr 107
|||
362 CTGGGGCCAGGGACCCAGGTCACCGTCTCCTCA 395
|||||
107 rTrpGlyGlnGlyThrLeuValThrValSerSer 118
```

seq\_name: sp\_human:Q96K68

seq\_documentation\_block:  
ID Q96K68 PRELIMINARY; PRT; 494 AA.

AC Q96K68;  
DT 01-DEC-2001 (TREMBLrel. 19, Created)  
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE CDNA FLJ14473 FIS, CLONE.MAMMA1001080, HIGHLY SIMILAR TO HOMO  
DE SAPIENS SNC73 PROTEIN (SNC73) MRNA.

OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]

RP SEQUENCE FROM N.A.  
RC TISSUE=MAMMARY GLAND;  
RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,  
RA Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,  
RA Wagatsuma M., Hosoiri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,  
RA Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,  
RA Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,  
RA Yamamoto J., Wakamatsu A., Nakamura Y., Nagahari K., Masuho Y.,  
RA Ninomiya K., Iwayanagi T.;  
RT "NEDO human cDNA sequencing project."  
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AK027379; BAB55072.1; --  
SQ SEQUENCE 494 AA; 53088 MW; 9A1D7AEB5AEE4C0E CRC64;

alignment\_scores:  
Quality: 364.00 Length: 130  
Ratio: 3.640 Gaps: 2  
Percent Similarity: 76.923 Percent Identity: 60.000

alignment\_block:  
US-09-742-690-1 x Q96K68 ..

Align seg 1/1 to: Q96K68 from: 1 to: 494

```
12 CAGGTGCAGTCGAGGAGTCAGGGGAGGATGGTGCAGGCTGGGGGCTC 61
:::|||||
20 GluValGlnLeuValGluSerGlyGlyGlyLeuValLysProGlyGlySe 36
|||||
62 TCTGAGACTCTCTGTGCAGCCTCGGGACCGCCGCCACCATGGTGCATGGTC 111
|||||
36 rLeuArgLeuSerCysAlaAlaSerGlyLeuSerPheSer.....T 50
|||||
112 ACTATGGTATGGGTGGTTCGCCAGGTTCCAGGGAAGGAGCGTGAGTTT 161
|||||
50 hrTyrAlaMetAsnTrpValArgGlnAlaProGlyLysGlyLeuGluTrp 66
|||||
162 GTCGCAGCTATTAGGTGGAGTGGTAAAGAGACATGGTATAAAGACTCCGT 211
|||||
67 ValSerSerIleSerSerArgSerAspTyrIleTyrTyrArgAspSerVa 83
|||||
212 GAAGGGCCGATTACCATCTCCAGAGATAACGCCAAGACTACGGTTTATC 261
|||||
83 lLysGlyArgPheThrIleSerArgAspAsnAlaLysAsnSerLeuTyrL 100
|||||
262 TGCAAAATGAACACCTGAAACCTGAAGATACGGCGCTTATTATTGTGCC 311
```

```
|||||
100 euGlnMetAsnSerLeuArgValAspThrAlaValTyrTyrCysAla 116
|||||
312 GCTCGACCGTCCGCGTGGATGATATTTCCCTGCGCGTTGGGTTTGACTA 361
:::
117 ArgAspSerCys.....AsnGlyAlaIleCysTyrGlyPheSerPr 130
|||
362 CTGGGGCCAGGGACCCAGGTCACCGTCTCCTCAGGATCT 401
|||||
130 oTrpGlyGlnGlyThrLeuValThrValSerSerAlaSer 143

seq_name: sp_human:Q9UL91

seq_documentation_block:  
ID Q9UL91 PRELIMINARY; PRT; 118 AA.  
AC Q9UL91;  
DT 01-MAY-2000 (TREMBLrel. 13, Created)  
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE MYOSIN-REACTIVE IMMUNOGLOBULIN HEAVY CHAIN VARIABLE REGION  
DE (FRAGMENT).  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=98277139; PubMed=9614934;  
RA Wu X., Liu B., Van der Merwe P.L., Kalis N.N., Berney S.M.,  
RA Young D.C.;  
RT "Myosin-reactive autoantibodies in rheumatic carditis and normal  
fetus."  
RL Clin. Immunol. Immunopathol. 87:184-192(1998).  
DR EMBL; AF035023; AAD56259.1; --  
DR HSSP; P01772; 2FB4.  
DR InterPro; IPR003006; Ig_MHC.  
DR InterPro; IPR003596; Ig_v.  
DR Pfam; PF00047; Ig; 1.  
DR SMART; SM00406; IGV; 1.  
FT NON_TER 1  
FT NON_TER 118  
SQ SEQUENCE 118 AA; 12843 MW; D0633949F2AC149D CRC64;
```

alignment\_scores:  
Quality: 362.00 Length: 127  
Ratio: 3.694 Gaps: 3  
Percent Similarity: 77.165 Percent Identity: 62.992

alignment\_block:  
US-09-742-690-1 x Q9UL91 ..

Align seg 1/1 to: Q9UL91 from: 1 to: 118

```
12 CAGGTGCAGTCGAGGAGTCAGGGGAGGATGGTGCAGGCTGGGGGCTC 61
:::|||||
1 GluValGlnLeuValGluSerGlyGlyGlyLeuValGlnProGlyGlySe 17
|||||
62 TCTGAGACTCTCTGTGCAGCCTCGGGACCGCCGCCACCATGGTGCATGGTC 111
|||||
17 rLeuArgLeuSerCysAlaAlaSerGlyPheThrPheSerSer..... 31
|||||
112 ACTATGGTATGGGTGGTTCGCCAGGTTCCAGGGAAGGAGCGTGAGTTT 161
|||||
32 ..TyrSerMetAsnTrpValArgGlnAlaProGlyLysGlyLeuGluTrp 47
|||||
162 GTCGCAGCTATTAGGTGGAGTGGTAAAGAGACATGGTATAAAGACTCCGT 211
|||||
48 ValSerTyrIleSerSerThrIleIleThrIleTyrTyrAlaAspSerVa 64
|||||
212 GAAGGGCCGATTACCATCTCCAGAGATAACGCCAAGACTACGGTTTATC 261
|||||
64 lLysGlyArgPheThrIleSerArgAspAsnAlaLysAsnSerLeuTyrL 81
```





```
117 ArgGlyGlyTyr.....PheAspVa 123
|||||
362 CTGGGGCCAGGGACCCAGGTCACCGTCTCCTCA 395
||||| ||||| ||||| ||||| |||||
123 lTrpGlyAlaGlyThrAlaValThrValSer 134
```

seq\_name: sp\_rodent:Q91Z07

```
seq_documentation_block:
ID Q91Z07 PRELIMINARY; PRT; 486 AA.
AC Q91Z07;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE HYPOTHETICAL 52.7 KDA PROTEIN.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DDBJ databases.
DR EMBL; BC010324; AAH10324.1; -.
KW Hypothetical protein.
SQ SEQUENCE 486 AA; 52682 MW; 4FEF835125DA870B CRC64;
```

alignment\_scores:
Quality: 346.50 Length: 136
Ratio: 3.397 Gaps: 5
Percent Similarity: 75.000 Percent Identity: 55.882

alignment\_block:
US-09-742-690-1 x Q91Z07 ..

Align seg 1/1 to: Q91Z07 from: 1 to: 486

```
12 CAGGTGCAGCTGCAGGAGTCAGGGGAGGATTGGTCAGGCTGGGGGCTC 61
:::||||| ||||| ||||| ||||| |||||
20 GluValHisLeuValGluSerGlyGlyLeuValLysProGlyGlySe 36
|||||
62 TCTGAGACTCTCCTGTGCAGCTCGGACGCGCCACAGTGGTCATGGTC 111
|||||
36 rLeuLysLeuSerCysValValSerGlyPheSerPheThrSer..... 50
112 ACTATGGTATGGCTGGTTCCGCCAGGTTCCAGGGAAGAGCGTGAGTTT 161
||| |||:|||| ||||| ||||| ||||| ||||| |||||
51 ..TyrAspMetSerTrpValArgGlnThrProGluArgArgLeuGluTrp 66
|||||
162 GTCCGACCTATTAGGTGGAGTGGTAAAGAGACATGGTATAAGACTCCGT 211
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
67 ValAlaAlaIleThr...SerGlyGlyAsnThrTyrTyrProAspAsnVa 82
212 GAAGGGCCGATTACCATCTCCAGAGATAACGCCAAGACTACGGTTTATC 261
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
82 lLysGlyArgPheThrValSerArgAspAsnAlaLysTyrThrLeuTyrL 99
262 TGCAAAATGAACAGCCTGAACCTGAAGATACGGCCGTTTATTATTGTGCC 311
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
99 euGlnMetSerSerLeuLysSerGluAspThrAlaMetTyrTyrCys... 114
312 GCTGACCGGTCGCGGTGGATGATATTTCCTGCGCGTT..... 350
:::||||| ||||| ||||| ||||| ||||| ||||| |||||
115 ValArgPro.....GluileProIleTyrTyrSe 125
351 .....GGGTTTGACTACTGGGCGCCAGGGACCCAGGTCACCGTCTCCT 393
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
125 rGlySerTyrPheAspSerTrpGlyGlnGlyThrThrIleThrValSerS 142
394 CAGGATCT 401
|| |||
```

```
142 erGluSer 144
seq_name: sp_rodent:Q91Z05
```

```
seq_documentation_block:
ID Q91Z05 PRELIMINARY; PRT; 473 AA.
AC Q91Z05;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE HYPOTHETICAL 51.9 KDA PROTEIN.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DDBJ databases.
DR EMBL; BC010327; AAH10327.1; -.
KW Hypothetical protein.
SQ SEQUENCE 473 AA; 51946 MW; CF625F008932AF12 CRC64;
```

alignment\_scores:
Quality: 342.00 Length: 128
Ratio: 3.600 Gaps: 3
Percent Similarity: 74.219 Percent Identity: 58.594

alignment\_block:
US-09-742-690-1 x Q91Z05 ..

Align seg 1/1 to: Q91Z05 from: 1 to: 473

```
12 CAGGTGCAGCTGCAGGAGTCAGGGGAGGATTGGTCAGGCTGGGGGCTC 61
:::||||| ||||| ||||| ||||| |||||
20 GluValGlnLeuValGluSerGlyGlyLeuValLysProGlyGlySe 36
|||||
62 TCTGAGACTCTCCTGTGCAGCTCGGACGCGCCACAGTGGTCATGGTC 111
| :::||||| ||||| ||||| ||||| |||||
36 rArgLysLeuSerCysAlaAlaSerGlyPheThrPheSer.....A 50
112 ACTATGGTATGGCTGGTTCCGCCAGGTTCCAGGGAAGAGCGTGAGTTT 161
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
50 sPtyrGlyMetHisTrpValArgGlnAlaProGluLysGlyLeuGluTrp 66
162 GTCCGACCTATTAGGTGGAGTGGTAAAGAGACATGGTATAAGACTCCGT 211
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
67 ValAlaTyrIleAsnSerGlySerThrThrIleTyrTyrAlaAspThrVa 83
212 GAAGGGCCGATTACCATCTCCAGAGATAACGCCAAGACTACGGTTTATC 261
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
83 lLysGlyArgPheThrIleSerArgAspAsnAlaLysAsnThrLeuPheL 100
262 TGCAAAATGAACAGCCTGAACCTGAAGATACGGCCGTTTATTATTGTGCC 311
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
100 euGlnMetThrSerLeuArgSerGluAspThrAlaMetTyrTyrCys... 115
312 GCTGACCGGTCGCGGTGGATGATATTTCCTGCGCGGTGGGTTGACTA 361
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
116 AlaArgGluLeuTrpLeuArgIle.....Aspty 126
362 CTGGGGCCAGGGACCCAGGTCACCGTCTCCTCA 395
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
126 rTrpGlyGlnGlyThrThrIleThrValSerSer 137
```

OM of: US-09-742-690-1 to: SwissProt\_40:\* out\_format : pfs  
Date: Aug 13, 2002 1:54 PM  
About: Results were produced by the GenCore software, version 4.5,  
Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:  
-MODEL=frame+n2p.model -DEV=xlp  
-Q=/cgn2\_1/USPTO\_spool/US09742690/runat\_12082002\_094916\_213/app\_query.fasta\_1.623  
-DB=SwissProt\_40 -QFMT=fastan -SUFFIX=rs -GAPOP=12.000  
-GAPEXT=4.000 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000  
-QGAPOP=4.500 -QGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500  
-FGAPOP=6.000 -FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500  
-DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=blossum62  
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR\_SCORE=pct  
-THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs  
-NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09742690.ecgn1\_1\_28 -NCPU=6 -ICPU=3 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -NO\_XLPXY -WAIT -THREADS=1

Search information block:  
Query: US-09-742-690-1  
Query length: 558  
Database: SwissProt\_40:\*  
Database sequences: 105224  
Database length: 38719550  
Search time (sec): 40.820000

score_list:	Sequence	Strid	Orig	zScore	EScore	Len	! Documentation
	SwissProt_40:HV3J_HUMAN	+	386.50	634.02	1.1e-27	121	! P01771 homo sapiens (human)
	SwissProt_40:HV3G_HUMAN	+	369.00	605.24	4.2e-26	122	! P01768 homo sapiens (human)
	SwissProt_40:HV3H_HUMAN	+	353.00	578.98	1.2e-24	122	! P01769 homo sapiens (human)
	SwissProt_40:HV3I_HUMAN	+	351.50	576.75	1.7e-24	119	! P01770 homo sapiens (human)
	SwissProt_40:HV3T_HUMAN	+	348.00	571.25	3.5e-24	116	! P01781 homo sapiens (human)
	SwissProt_40:HV3U_HUMAN	+	347.00	569.29	4.3e-24	120	! P01782 homo sapiens (human)
	SwissProt_40:HV3C_HUMAN	+	346.50	568.71	4.8e-24	117	! P01764 homo sapiens (human)
	SwissProt_40:HV3K_HUMAN	+	346.00	567.20	5.4e-24	126	! P01772 homo sapiens (human)
	SwissProt_40:HV01_CANFA	+	344.00	564.84	8.0e-24	114	! P01784 canis familiaris (dog)
	SwissProt_40:HV3F_HUMAN	+	343.50	563.94	8.9e-24	115	! P01767 homo sapiens (human)
	SwissProt_40:HV3E_HUMAN	+	343.50	563.55	9.0e-24	120	! P01766 homo sapiens (human)
	SwissProt_40:HV3A_HUMAN	+	341.00	559.30	1.5e-23	122	! P01762 homo sapiens (human)
	SwissProt_40:HV33_MOUSE	+	340.50	559.02	1.7e-23	115	! P01802 mus musculus (mouse)
	SwissProt_40:HV3L_HUMAN	+	340.50	558.70	1.7e-23	119	! P01773 homo sapiens (human)
	SwissProt_40:HV32_MOUSE	+	337.50	554.10	3.2e-23	115	! P01801 mus musculus (mouse)
	SwissProt_40:HV30_MOUSE	+	336.50	552.62	3.9e-23	113	! P01799 mus musculus (mouse)
	SwissProt_40:HV31_MOUSE	+	333.50	547.70	7.3e-23	113	! P01800 mus musculus (mouse)
	SwissProt_40:HV02_CANFA	+	333.50	547.38	7.3e-23	117	! P01785 canis familiaris (dog)
	SwissProt_40:HV27_MOUSE	+	331.50	544.41	1.1e-22	113	! P01796 mus musculus (mouse)
	SwissProt_40:HV3D_HUMAN	+	331.50	544.25	1.1e-22	115	! P01765 homo sapiens (human)
	SwissProt_40:HV16_MOUSE	+	331.00	541.89	1.3e-22	136	! P01783 mus musculus (mouse)
	SwissProt_40:HV35_MOUSE	+	330.00	542.12	1.5e-22	111	! P01804 mus musculus (mouse)
	SwissProt_40:HV3B_HUMAN	+	329.00	540.23	1.9e-22	114	! P01763 homo sapiens (human)
	SwissProt_40:HV05_CARAU	+	328.00	538.43	2.3e-22	116	! P19181 carassius auratus (goldfish)
	SwissProt_40:HV55_MOUSE	+	326.50	535.89	3.2e-22	117	! P18526 mus musculus (mouse)
	SwissProt_40:HV28_MOUSE	+	325.50	534.57	3.9e-22	113	! P01797 mus musculus (mouse)
	SwissProt_40:HV29_MOUSE	+	321.50	528.01	9.1e-22	113	! P01798 mus musculus (mouse)
	SwissProt_40:HV53_MOUSE	+	319.50	524.41	1.4e-21	117	! P18524 mus musculus (mouse)
	SwissProt_40:HV41_MOUSE	+	314.50	516.20	4.0e-21	117	! P01811 mus musculus (mouse)
	SwissProt_40:HV57_MOUSE	+	313.50	516.19	4.8e-21	98	! P18528 mus musculus (mouse)
	SwissProt_40:HV54_MOUSE	+	313.50	514.56	4.9e-21	117	! P18525 mus musculus (mouse)
	SwissProt_40:HV25_MOUSE	+	313.50	514.10	5.0e-21	123	! P01794 mus musculus (mouse)
	SwissProt_40:HV3N_HUMAN	+	312.00	511.94	6.8e-21	119	! P01775 homo sapiens (human)
	SwissProt_40:HV20_MOUSE	+	312.00	511.71	6.8e-21	122	! P01789 mus musculus (mouse)
	SwissProt_40:HV22_MOUSE	+	311.50	510.82	7.6e-21	123	! P01791 mus musculus (mouse)
	SwissProt_40:HV3M_HUMAN	+	311.00	510.30	8.4e-21	119	! P01774 homo sapiens (human)
	SwissProt_40:HV56_MOUSE	+	308.00	507.27	1.5e-20	97	! P18527 mus musculus (mouse)
	SwissProt_40:HV19_MOUSE	+	307.50	504.26	1.8e-20	123	! P01788 mus musculus (mouse)
	SwissProt_40:HV21_MOUSE	+	307.00	503.51	2.0e-20	122	! P01790 mus musculus (mouse)
	SwissProt_40:HV39_MOUSE	+	306.50	503.00	2.2e-20	118	! P01809 mus musculus (mouse)

SwissProt\_40:HV01\_RAT + 306.50 501.29 2.2e-20 142 ! P01805 rattus norvegicus (r  
SwissProt\_40:HV38\_MOUSE + 306.00 502.10 2.4e-20 119 ! P01808 mus musculus (mous  
SwissProt\_40:HV3P\_HUMAN + 306.00 502.10 2.4e-20 119 ! P01777 homo sapiens (huma  
SwissProt\_40:HV18\_MOUSE + 305.50 500.97 2.7e-20 123 ! P01787 mus musculus (mous  
SwissProt\_40:HV34\_MOUSE + 304.50 500.11 3.3e-20 113 ! P01803 mus musculus (mous  
  
seq\_name: SwissProt\_40:HV3J\_HUMAN  
  
seq\_documentation\_block:  
ID HV3J\_HUMAN STANDARD; PRT; 121 AA.  
AC P01771;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE Ig heavy chain V-III region HIL.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=79124695; PubMed=420800;  
RA Chiu Y.-Y.H., Lopez de Castro J.A., Poljak R.J.;  
RT "Amino acid sequence of the VH region of human myeloma  
cryoimmunoglobulin IgG Hil.";  
RL Biochemistry 18:553-560(1979).  
CC -1- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM AN IGG1 MYELOMA  
CC PROTEIN.  
DR PIR; A02054; G1HUHL.  
DR HSSP; P01772; 2FB4.  
DR InterPro; IPR003006; Iq\_MHC.  
DR InterPro; IPR003596; Iq\_v.  
DR Pfam; PF00047; ig; 1.  
DR SMART; SM00406; IGV; 1.  
DR KW Immunoglobulin V region.  
FT MOD\_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.  
FT NON\_TER 121 121  
SQ SEQUENCE 121 AA; 13566 MW; 480FC53610EF5DAB CRC64;  
  
alignment\_scores:  
Quality: 386.50 Length: 128  
Ratio: 3.865 Gaps: 2  
Percent Similarity: 78.125 Percent Identity: 61.719  
  
alignment\_block:  
US-09-742-690-1 x HV3J\_HUMAN ..  
  
Align seg 1/1 to: HV3J\_HUMAN from: 1 to: 121  
  
12 CAGGTGCAGCTGCAGGAGTCAAGGGGAGGATTGGTGCAGGCTGGGGCTC 61  
|||||::||| :::::|||||:::||||| ||| |||  
1 GlnValLysLeuValGlnAlaGlyGlyValValGlnProGlyArgse 17  
  
62 TCTGAGACTCTCCTGTGCAGCTCGGGACGCCACCACTGTCATGGTC 111  
|||||::||| ||| |||||:::||||| ::| |||  
17 rLeuArgLeuSerCysIleAlaSerGlyPheThrPheSer.....A 31  
  
112 ACTATGGTATGGCTGGTCCGCCAGGTTCCAGGAAGGAGCGTGAGTTT 161  
:::||||| ||| |||||:::||||| |||::|  
31 snTyGlyMetHisTrpValArgGlnAlaProGlyLysGlyLeuGluTrp 47  
  
162 GTCGCAGCTATTAGTGGAGTGGTAAAGAGACATGGTATAAGACTCCGT 211  
|||||::||| :::::|||||:::||||| ||| ||| ||| |||  
48 ValAlaValIleTrpTyArgAsnGlySerArgThrTyTrpGlyAspSerVa 64  
  
212 GAAGGGCGGATTCCACCATCTCCAGAGATAACGCCAAGACTACGGTTATC 261  
|||||::||| |||||:::||||| |||||::||| |||::|||  
64 lLysGlyArgPheThrIleSerArgAspAsnSerLysArgThrLeuTyrm 81  
  
262 TGCAATGAACAGCCCTGAAACCTGAAGATACGGCCGTTTATTATTGTC 311  
:::||||| |||||:::||||| |||||:::||||| |||||  
81 et\*\*\*MetAsnSerLeuArgThrGluAspThrAlaValTyTrpCysala 97





64 lLysGlyArgPheThrIleSerArg\*\*\*\*\*SerLys\*\*\*ThrMetTyrL 81  
262 TGC AAAATGAACAGCCCTGAAACCTGAAGATACGGCGCTTTATTATTGTGCC 311  
||||| ||||| :||| :||||| ||||| |||||  
81 euGluMetAsnSerLeuArgAlaGluAsnThrAlaValTyrCysAla 97  
312 GCTCGACCGGTCCGGCGTGATGATATTTCCCTGCCGGTTGGGTTTGACTA 361  
::: :::: :||| |||||  
98 ArgSerGlyIleAlaLeuGlySerValAla.....GlyThrAspTy 111  
362 CTGGGGCCAGGGGACCAGGTCACCGTCTCCTCA 395  
||||| :||| ||||| |||||  
111 rTrpGly\*\*\*GlyThrLeuValThrIleSerSer 122

seq\_name: SwissProt\_40:HV3I\_HUMAN

```
seq_documentation_block:
ID HV3I HUMAN STANDARD:
PRT: 119 AA:
```

AC P01770;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE Ig heavy chain V-III region NIE.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=77070269; PubMed=826475;  
RA Ponstingl H., Hilschmann N.;  
RT "The rule of antibody structure. The primary structure of a  
RT monoclonal IgG1 immunoglobulin (myeloma protein Nie). III. The  
RT chymotryptic peptides of the H-chain, alignment of the tryptic  
RT peptides and discussion of the complete structure.";  
RL Hoppe-Seyler's Z. Physiol. Chem. 357:1571-1604(1976)

alignment_scores:		
Quality:	351.50	Length: 128
Ratio:	3.515	Gaps: 3
Percent Similarity:	78.125	Percent Identity: 58.594

alignment\_block:

US-09-742-690-1 X HV3I HUMAN

Align seq 1/1 to: HV31 HUMAN from: 1 to: 119

12 CAGGTGCAGCTGCAGGAGTCAGGGGGAGGATTGGTGCAGGCTGGGGGGCTC 61  
 |||||  
 ::|||  
 1 GlnValGlnLeuValGlnSerGlyGlyValValGlnProGlyArgSer 17

[illegible]seq name: SwissProt 40:HV3T\_HUMAN

seq\_documentation\_block:  
ID HV37T HUMAN STANDARD: PRT: 116 AA.

ID AC P01781;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE Ig heavy chain V-III region GAL.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=75059123; PubMed=4803843;  
 RA Watanabe S., Barnikol H.U., Horn J., Bertram J., Hilschmann N.;  
 RT "The primary structure of a monoclonal IgM-immunoglobulin  
 RT (macroglobulin Gal.), II: the amino acid sequence of the H-chain (mu-  
 RT type), subgroup H III. Architecture of the complete IgM-molecule.";  
 RT Horne-Severin's % physiol. Chem 354:1505-1509(1973).

alignment_scores:		
Quality:	348.00	Length: 128
Ratio:	3.824	Gaps: 2

alignment\_block:

```
US-09-742-690-1 x HV3U_HUMAN ..
Align seg 1/1 to: HV3U_HUMAN from: 1 to: 116
12 CAGGTGCAGCTGCAGGAGTCAGGGGGAGGATTGGTGCAGGCTGGGGGCTC 61
:::|||||  |||||  |||||  |||||  |||||  |||||  |||||  ||
1 GluValGlnLeuValGluSerGlyGlyAspLeuValGlnProGlyArgSe 17
62 TCTGAGACTCTCCTGTGTCAGCCTCGGACGGCCACCACCGTGGTCATGGTC 111
|||||  |||||  |||||  |||||  |||||  |||||  |||||  ||
17 rLeuArgLeuSerCysAlaAlaSerGly.....Phe***Phe*** 31
112 ACTATGGTATGGGCTGTTCCGCCAGGTTCCAGGGAAGAGCGTGAGTTT 161
:::  |||||  |||  |||||  :::  |||||  |||||  |||||  ||
31 **LeuGlyMetThrTrpValArgGlnAlaProGlyLysGlyLeuGluTrp 47
162 GTCGACGCTATTAGTGGAGTGGTAAAGAGACATGGTATAAAGACTCCGT 211
|||||  |||||  :::  |||||  :::  |||||  |||||  |||||
48 ValAlaAsnIleLys****GlySer*****TyrValAspServa 64
212 GAAGGCCGATTACCATCTCCAGAGATAACGCCAAGACTACGGTTTATC 261
|||||  |||||  |||||  |||||  |||||  |||||  |||||  ||
64 lLysGlyArgPheThrIleSerArgAspAsnAlaLysAsnSerLeuTyrL 81
262 TGCAAAATGAACAGCCTGAACCTGAAGATACGGCCGTTTATTATTGTGCC 311
|||||  |||||  |||||  |||||  |||||  |||||  |||||  ||
81 euGlnMetAsnSerLeuArgValGluAspThrAlaLeuTyrTyrCysAla 97
312 GCTCGACCGGTCGGCGTGGATGATATTCCCTGCCGTTGGGTTTGACTA 361
|||  |||||
98 ArgGly.....TrpGlyGlyGlyAspTy 105
362 CTGGGGCCAGGGACCCAGGTCACCGTCTCCTCA 395
|||||  |||||  |||||  |||||  |||||  |||||  |||||  ||
105 rTrpGlyGlnGlyThrLeuValThrValSerThr 116
seq_name: SwissProt_40:HV3U_HUMAN
seq_documentation_block:
ID HV3U_HUMAN STANDARD; PRT; 120 AA.
AC P01782;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig heavy chain V-III region DOB.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE.
RX MEDLINE=80020921; Pubmed=114209;
RA Steiner L.A., Garcia Pardo A., Margolies M.N.;
RT "Amino acid sequence of the heavy-chain variable region of the
RT crystallizable human myeloma protein Dob.";
RL Biochemistry 18:4068-4080(1979).
RN [2]
RP CRYSTALLIZATION.
RX MEDLINE=80020920; Pubmed=114208;
RA Steiner L.A., Lopes A.D.;
RT "The crystallizable human myeloma protein Dob has a hinge-region
RT deletion.";
RL Biochemistry 18:4054-4067(1979).
CC -!- MISCELLANEOUS: THIS GAMMA-1 MYELOMA PROTEIN HAS A DELETION IN THE
CC HINGE REGION. THERE ARE NO LIGHT-HEAVY OR INTER-HEAVY CHAIN
CC DISULFIDE BONDS.
DR PIR; A02065; GLHUBD.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_V.
DR Pfam; PF00047; ig; 1.
DR SMART; SM00406; IGV; 1.
KW Immunoglobulin V region.
FT NON_TER 120 120
```

```
SQ SEQUENCE 120 AA; 13440 MW; 880DDE307C4B2627 CRC64;
alignment_scores:
Quality: 347.00 Length: 128
Ratio: 3.731 Gaps: 2
Percent Similarity: 72.656 Percent Identity: 58.594
alignment_block:
US-09-742-690-1 x HV3U_HUMAN ..
Align seg 1/1 to: HV3U_HUMAN from: 1 to: 120
12 CAGGTGCAGCTGCAGGAGTCAGGGGGAGGATTGGTGCAGGCTGGGGGCTC 61
:::|||||  |||||  |||||  |||||  |||||  |||||  |||||  ||
1 GluValGlnLeuValGluSerGlyGlyAspLeuValGlnProGlyArgSe 17
62 TCTGAGACTCTCCTGTGTCAGCCTCGGACGGCCACCACCGTGGTCATGGTC 111
|||||  |||||  |||||  |||||  |||||  |||||  |||||  ||
17 rLeuArgLeuSerCysAlaAlaSerGly.....PheAsnPheHisG 31
112 ACTATGGTATGGGCTGTTCCGCCAGGTTCCAGGGAAGAGCGTGAGTTT 161
:::  |||||  |||||  |||||  |||||  |||||  |||||  ||
31 lutyAsnMetHisTrpLeuArgGlnGlyProGlyLysGlyProGluTrp 47
162 GTCGACGCTATTAGTGGAGTGGTAAAGAGACATGGTATAAAGACTCCGT 211
|||||  |||||  |||||  |||||  |||||  |||||  |||||  ||
48 ValSerThrIleThrTrpAsnGlyGlySerValLeuTyrAlaAspServa 64
212 GAAGGCCGATTACCATCTCCAGAGATAACGCCAAGACTACGGTTTATC 261
|||||  |||||  |||||  |||||  |||||  |||||  |||||  ||
64 lLysGlyArgPheAlaIleSerArgAspAsnAlaGlnLysThrLeuTyrL 81
262 TGCAAAATGAACAGCCTGAACCTGAAGATACGGCCGTTTATTATTGTGCC 311
|||||  |||||  |||||  |||||  |||||  |||||  |||||  ||
81 euGlnLeuAsnIleLeuArgProGluAspThrAlaPheTyrTyrCysAla 97
312 GCTCGACCGGTCGGCGTGGATGATATTCCCTGCCGTTGGGTTTGACTA 361
|||  |||||
98 LysGlyTyrIleTrpAsnGlyAsn.....TrpPheAspSe 109
362 CTGGGGCCAGGGACCCAGGTCACCGTCTCCTCA 395
|||||  |||||  |||||  |||||  |||||  |||||  |||||  ||
109 rTrpGlyGlnGlyThrLeuValThrValSerSer 120
seq_name: SwissProt_40:HV3C_HUMAN
seq_documentation_block:
ID HV3C_HUMAN STANDARD; PRT; 117 AA.
AC P01764;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig heavy chain V-III region VH26 precursor.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=81101090; Pubmed=6450418;
RA Matthysens G., Rabbitts T.H.;
RT "Structure and multiplicity of genes for the human immunoglobulin
RT heavy chain variable region.";
RL Proc. Natl. Acad. Sci. U.S.A. 77:6561-6565(1980).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/
CC or send an email to license@isb-sib.ch).
```

```
CC -----
DR EMBL; J00236; AAA53516.1; -.
DR EMBL; M35415; AAA58735.1; -.
DR PIR; A02047; H3HU26.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; ig; 1.
DR SMART; SM00406; Igv; 1.
KW Immunoglobulin V region; Signal.
FT SIGNAL 1 19
FT CHAIN 20 117 IG HEAVY CHAIN V-III REGION VH26.
FT NON_TER 117 117
SQ SEQUENCE 117 AA; 12582 MW; E826733F1A3CB0F1 CRC64;

alignment_scores:
  Quality: 346.50      Length: 100
  Ratio: 4.076        Gaps: 1
  Percent Similarity: 85.000      Percent Identity: 71.000

alignment_block:
US-09-742-690-1 x HV3C_HUMAN ..

Align seg 1/1 to: HV3C_HUMAN from: 1 to: 117

12 CAGGTGCAGCTCAGGAGTCAAGGGGAGGATGGTGCAGGCTGGGGGCTC 61
   ::::::::::::::::::::
20 GluValGlnLeuLeuGluSerGlyGlyLeuValGlnProGlyGlySe 36
   ::::::::::::::::::::
62 TCTGAGACTCTCCTGTGTCAGCCCTCGGAGCGGCCACCAGTGGTCATGGTC 111
   ::::::::::::::::::::
36 rLeuArgLeuSerCysAlaAlaSerGlyPheThrPheSerSer..... 50
   ::::::::::::::::::::
112 ACTATGGTATGGGCTGGTTCGCCAGGTTCCAGGGAAGGAGCGTGAGTTT 161
   ::::::::::::::::::::
51 ..TyrAlaMetSerTrpValArgGlnAlaProGlyLysGlyLeuGluTrp 66
   ::::::::::::::::::::
162 GTCGACGCTATTAGTGGAGTGGTAAAGAGACATGGTATAAAGACTCCGT 211
   ::::::::::::::::::::
67 ValSerAlaIleSerGlySerGlyGlySerThrTyrGlyAspSerVa 83
   ::::::::::::::::::::
212 GAAGGGCCGATTCACCATCTCCAGAGATAACGCCAAGACTACGGTTATC 261
   ::::::::::::::::::::
83 lLysGlyArgPheThrIleSerArgAspAsnSerLysAsnThrLeuTyrL 100
   ::::::::::::::::::::
262 TGCAAAATGAACAGCCTGAAACCTGAAGATACGGCCGTTTATTATTGTC 311
   ::::::::::::::::::::
100 euGlnMetAsnSerLeuArgAlaGluAspThrAlaValTyrTyrCysAla 116
   ::::::::::::::::::::

seq_name: SwissProt_40:HV3K_HUMAN

seq_documentation_block:
ID HV3K_HUMAN STANDARD; PRT; 126 AA.
AC P01772;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig heavy chain V-III region KOL.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE, AND DISULFIDE BONDS.
RX MEDLINE=83289131; PubMed=6884994;
RA Schmidt W.E., Jung H.-D., Palm W., Hilschmann N.;
RT "Three-dimensional structure determination of antibodies. Primary
RT structure of crystallized monoclonal immunoglobulin IgG1 KOL, I.";
RL Hoppe-Seyler's Z. Physiol. Chem. 364:713-747(1983).
RN [2]
RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS).
RX MEDLINE=81072295; PubMed=7441755;
RA Marquart M., Deisenhofer J., Huber R., Palm W.;
```

```
RT "Crystallographic refinement and atomic models of the intact
RT immunoglobulin molecule Kol and its antigen-binding fragment at 3.0 A
RT and 1.0-A resolution.";
RL J. Mol. Biol. 141:369-391(1980).
DR PIR; A02055; GIHUKL.
DR PDB; 2FB4; 12-JUL-89.
DR PDB; 2IG2; 12-JUL-89.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; ig; 1.
DR SMART; SM00406; Igv; 1.
KW Immunoglobulin V region; 3D-structure.
FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
FT DISULFID 22 96
FT DISULFID 105 110
FT STRAND 3 7
FT STRAND 11 12
FT TURN 14 15
FT STRAND 18 25
FT HELIX 29 31
FT STRAND 34 39
FT TURN 41 42
FT STRAND 46 51
FT TURN 53 54
FT STRAND 58 60
FT TURN 62 67
FT STRAND 68 73
FT TURN 74 77
FT STRAND 78 83
FT HELIX 88 90
FT STRAND 92 99
FT STRAND 106 106
FT TURN 107 108
FT STRAND 109 109
FT STRAND 113 113
FT STRAND 120 124
FT NON_TER 126 126
SQ SEQUENCE 126 AA; 13718 MW; E4D71B52B16F8776 CRC64;

alignment_scores:
  Quality: 346.00      Length: 137
  Ratio: 3.604        Gaps: 3
  Percent Similarity: 70.073      Percent Identity: 56.204

alignment_block:
US-09-742-690-1 x HV3K_HUMAN ..

Align seg 1/1 to: HV3K_HUMAN from: 1 to: 126

12 CAGGTGCAGCTGCAGGAGTCAAGGGGAGGATGGTGCAGGCTGGGGGCTC 61
   ::::::::::::::::::::
1 GlnValGlnLeuValGluSerGlyGlyValValGlnProGlyArgSe 17
   ::::::::::::::::::::
62 TCTGAGACTCTCCTGTGTCAGCCCTCGGAGCGGCCACCAGTGGTCATGGTC 111
   ::::::::::::::::::::
17 rLeuArgLeuSerCysSerSerSerGlyPheIlePheSerSer..... 31
   ::::::::::::::::::::
112 ACTATGGTATGGGCTGGTTCGCCAGGTTCCAGGGAAGGAGCGTGAGTTT 161
   ::::::::::::::::::::
32 ..TyrAlaMetTyrTrpValArgGlnAlaProGlyLysGlyLeuGluTrp 47
   ::::::::::::::::::::
162 GTCGACGCTATTAGTGGAGTGGTAAAGAGACATGGTATAAAGACTCCGT 211
   ::::::::::::::::::::
48 ValAlaIleIleTrpAspAspGlySerAspGlnHisTyrAlaAspSera 64
   ::::::::::::::::::::
212 GAAGGGCCGATTCACCATCTCCAGAGATAACGCCAAGACTACGGTTATC 261
   ::::::::::::::::::::
64 lLysGlyArgPheThrIleSerArgAspAsnSerLysAsnThrLeuPheL 81
   ::::::::::::::::::::
262 TGCAAAATGAACAGCCTGAAACCTGAAGATACGGCCGTTTATTATTGTC 311
   ::::::::::::::::::::
81 euGlnMetAspSerLeuArgProGluAspThrGlyValTyrPheCysAla 97
   ::::::::::::::::::::
```

312 GCTGACCGGTCCTGGTGATGATATTTCCCTGCGCGTGGGTTT..... 356  
                  :::  
98 ArgAspGlyGlyHis.....GlyPheCysSe 106  
                  :::  
357 .....GACTACTGGGGCCAGGGGACCCAGGTCA 384  
                  |||||  
106 rSerAlaSerCysPheGlyProAspTyrTrpGlyGlnGlyThrProValT 123  
                  |||||

385 CCGTCTCCTCA 395  
      |||||  
123 hrValSerSer 126

seq\_name: SwissProt\_40:HV01\_CANFA

seq\_documentation\_block:

ID HV01\_CANFA STANDARD; PRT; 114 AA.

AC P01784;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE Ig heavy chain V region GOM.  
OS Canis familiaris (Dog).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.  
OX NCBI\_TaxID=9615;  
RN [1]  
RP SEQUENCE.

RX MEDLINE-77242268; PubMed-407924;  
RA Wasserman R.L., Capra J.D.;  
RT "Primary structure of the variable regions of two canine  
immunoglobulin heavy chains.";  
RL Biochemistry 16:3160-3168(1977).  
CC -I- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM A MYELOMA PROTEIN.  
DR PIR; A02067; AVDGGM.  
DR InterPro; IPR003006; Ig\_MHC.  
DR InterPro; IPR003596; Ig\_v.  
DR Pfam; PF00047; ig; 1.  
DR SMART; SM00406; Igv; 1.  
KW Immunoglobulin V region.  
FT NON\_TER 114 114  
SQ SEQUENCE 114 AA; 12430 MW; B1D4745D2C4E13C4 CRC64;

alignment\_scores:  
  Quality: 344.00      Length: 128  
  Ratio: 3.780        Gaps: 2  
Percent Similarity: 71.094    Percent Identity: 58.594

alignment\_block:

US-09-742-690-1 x HV01\_CANFA ..

Align seg 1/1 to: HV01\_CANFA from: 1 to: 114

12 CAGGTGCAGTCGACGAGTCAGGGGAGGATTGGTCAGGCTGGGGGCTC 61  
  :::|||||    |||||    |||||    |||||    |||||  
1 GluValGlnLeuValGluSerGlyGlyAspLeuValLysProGlyGlySe 17  
62 TCTGAGACTCTCCTGTGCAGCCTCGGGACGCGCCACCAGTCGTCATGGTC 111  
  |||||    |||||    |||||    |||||    |||||  
17 rLeuArgLeuSerCysValAlaSerGlyIleThrPheSerGly..... 31  
112 ACTATGGTATGGGCTGTTCCGCCAGGTTCCAGGGAAGGAGCGTGAGTTT 161  
  |||    |||    |||||    |||||    |||||  
32 ..TyrAspMetGlnTrpValArgGlnAlaProGlyLysGlyLeuGlnLys 47  
162 GTCGCAGCTATTAGTGGAGTGGTAAAGACACATGGTATAAAGACTCCGT 211  
  |||||    :::::    :::::    |||    |||||  
48 ValAlaTyrPheAsnAspAlaLeuSerAlaGlnGlyTyrAlaAspAlaVa 64  
212 GAAGGGCCGATTCCACCATCTCCAGAGATAACGCCACACTACGGTTTATC 261  
  |||||    |||||    |||||    |||||    |||||    |||||  
64 lLysGlyArgPheThrIleSerLysAspAsnAlaLysAspSerLeuTyrL 81

262 TCAAAATGAACAGCCTGAAACCTGAAGATACGGCGTTTATTATTGTGCC 311  
  |||||    |||||    |||||    |||||    |||||    |||||  
81 euGlnMetAsnSerLeuArgAlaGluAspThrAlaValTyrTyrCysAla 97  
312 GCTCGACCGGTCCTCGCGTGATGATATTTCCCTGCCGTTGGGTTGACTA 361  
  |||    |||    |||||  
98 .....ProTrpGlnPheGluTy 103  
362 CTGGGGCCAGGGGACCCAGGTCACCGTCTCCTCA 395  
  |||||    |||||    |||||    |||||  
103 rTrpGlyGlnGlyThrLeuValThrValSerSer 114

seq\_name: SwissProt\_40:HV3F\_HUMAN

seq\_documentation\_block:

ID HV3F\_HUMAN STANDARD; PRT; 115 AA.

AC P01767;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE Ig heavy chain V-III region BUT.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE-78137069; PubMed-416441;  
RA Torano A., Putnam F.W.;  
RT "Complete amino acid sequence of the alpha 2 heavy chain of a human  
  -IgA2 immunoglobulin of the A2m (2) allotype.";  
RL Proc. Natl. Acad. Sci. U.S.A. 75:966-969(1978).  
CC -I- MISCELLANEOUS: THE SEQUENCE OF THE ALPHA-2, A2M(2) ALLOTYPE, C  
  REGION OF THIS MYELOMA PROTEIN IS ALSO GIVEN.  
DR PIR; A02050; A2HUBU.  
DR HSSP; P01789; IMCP.  
DR InterPro; IPR003006; Ig\_MHC.  
DR InterPro; IPR003596; Ig\_v.  
DR Pfam; PF00047; ig; 1.  
DR SMART; SM00406; Igv; 1.  
KW Immunoglobulin V region.  
FT NON\_TER 115 115  
SQ SEQUENCE 115 AA; 12379 MW; 208876A7DF52DCF4 CRC64;

alignment\_scores:  
  Quality: 343.50      Length: 128  
  Ratio: 3.578        Gaps: 4  
Percent Similarity: 75.000    Percent Identity: 58.594

alignment\_block:

US-09-742-690-1 x HV3F\_HUMAN ..

Align seg 1/1 to: HV3F\_HUMAN from: 1 to: 115

12 CAGGTGCAGTCGACGAGTCAGGGGAGGATTGGTCAGGCTGGGGGCTC 61  
  :::|||||    |||||    |||||    |||||    |||||  
1 GluValGlnLeuValGluThrGlyGlyGlyLeuIleGlnProGlyGlySe 17  
62 TCTGAGACTCTCCTGTGCAGCCTCGGGACGCGCCACCAGTCGTCATGGTC 111  
  |||||    |||||    |||||    |||||    |||||  
17 rLeuArgLeuSerCysAlaAlaSerGlyPheThrValSer\*\*His.... 32  
112 ACTATGGTATGGGCTGTTCCGCCAGGTTCCAGGGAAGGAGCGTGAGTTT 161  
  :::|||||    |||||    |||||    |||||    |||||  
33 .....SerMetSerTrpValArgGlnAlaProGlyLysAlaLeu\*\*Trp 47  
162 GTCGCAGCTATTAGGTGGAGTGGTAAAGAGACATGGTATAAAGACTCCGT 211  
  |||||    |||||    :::::    |||    |||||  
48 ValSerAlaIle...TyrArgGlyGlyThrThrTyrTyrAlaAspSerVa 63  
212 GAAGGGCCGATTCCACCATCTCCAGAGATAACGCCAAGACTACGGTTTATC 261



```
|||||
63 llySGlyArgPheThrIleSerArgAspSerArg**ThrValTyrL 80
262 TGCAAATGAACAGCCTGAACCTGAAGATACGGCGCTTATTATTGTGCC 311
|||||
80 euGlnMet***SerLeuArgAlaGluAspThrAlaValTyrTyrCysAla 96
312 GCTCGACCGGTCCGCGTGGATGATATTCCCTCGCGGTGGGTTTGACTA 361
|||||
97 Arg.....AspLeuAlaAlaAlaArgLeuPhe..... 105
362 CTGGGGCCAGGGGACCCAGCTCACCGTCTCCTCTCA 395
|||||
106 ....GlyLysGlyThrThrValThrValSerSer 115
```

seq\_name: SwissProt\_40:HV3E\_HUMAN

seq\_documentation\_block:  
ID HV3E\_HUMAN STANDARD; PRT; 120 AA.

AC P01766;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE Ig heavy chain V-III region BRO.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheraia; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=77117674; PubMed=65324;  
RA Capra J.D., Hopper J.E.;  
RT "Comparative studies on monotypic IgM lambda and IgG kappa from an individual patient. III. The complete amino acid sequence of the VH region of the IgM paraprotein.";  
RT Immunohistochemistry 13:995-999(1976).  
RL -I- MISCELLANEOUS: THIS CHAIN WAS OBTAINED FROM IGM ISOLATED FROM THE SERUM OF A PATIENT WITH MALIGNANT LYMPHOMA OF THE WALDENSTROM TYPE.

DR PIR; A02049; M3HUBW.  
DR InterPro; IPR003006; Ig\_MHC.  
DR InterPro; IPR003596; Ig\_v.  
DR Pfam; PF00047; ig; 1.  
DR SMART; SM00406; IGV; 1.  
KW Immunoglobulin V region.  
FT NON\_TER 120 120  
SQ SEQUENCE 120 AA; 13227 MW; D3F0428F7C2E6410 CRC64;

alignment\_scores:  
Quality: 343.50 Length: 128  
Ratio: 3.616 Gaps: 4  
Percent Similarity: 74.219 Percent Identity: 60.156

alignment\_block:

US-09-742-690-1 x HV3E\_HUMAN ..

Align seg 1/1 to: HV3E\_HUMAN from: 1 to: 120

```
12 CAGGTGCAGCTGCAGGAGTCAAGGGGAGGATTGGTGCAGGCTGGGGGCTC 61
:::|||||
1 GluValGlnLeuValGluSerGlyGlyGlyLeuValGlnProGlyGlyse 17
62 TCTGAGACTCTCTGTGCAGCCTCGGGACGCGCCACACAGTGGTCATGGTC 111
|||||
17 rLeuArgLeuSerCysAlaAlaSerGlyPheThrPheSer.....T 31
112 ACTATGGTATGGGCTGTCCGCCAGGTTCCAGGGAAGGAGCGTGAGTTT 161
:::|||||
31 yTyrAsnMetAsnTrpValArgGlnValThrGlyLysGlyLeuGluTrp 47
162 GTCGCAGCTATTAGGTGGAGTGGTAAAGAGACATGGTATAAAGACTCCGT 211
|||||
```

```
48 ValSerAlaIleGlyThrAlaGly...AspGlnTyrTyrAlaAspSerVa 63
212 GAAGGGCCGATTCCACCATCTCCAGAGATAACCCCAAGACTACGGTTTATC 261
|||||
63 llySGlyArgPheThrIleSerArgAsnAspSerLysAsnThrLeuTyrL 80
262 TGCAAATGAACAGCCTGAACCTGAAGATACGGCGCTTATTATTGTGCC 311
|||||
80 euAsnMetAsnSerLeuArgAlaGluAspThrAlaValTyrTyrCysAla 96
312 GCTCGACCGGTCCGCGTGGATGATATTCCCTCGCGGTGGGTTTGACTA 361
|||||
97 ArgSerProVal.....SerLeuValAspGlyTrpLeuTy 108
362 C.....TGGGGCCAGGGGACC 377
108 rTyrTyrTyrGlySerValTrpGlyGlnGlyThr 119
```

seq\_name: SwissProt\_40:HV3A\_HUMAN

seq\_documentation\_block:  
ID HV3A\_HUMAN STANDARD; PRT; 122 AA.

AC P01762;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE Ig heavy chain V-III region TRO.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheraia; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE (MYELOMA PROTEIN TRO).  
RX MEDLINE=76023781; PubMed=809331;  
RA Kratzin H., Altevogt P., Ruban E., Kortt A., Starosik K., Hilschmann N.;  
RT "The primary structure of a monoclonal IgA-immunoglobulin (IgA Tro.), II. The amino acid sequence of the H-chain, alpha-type, subgroup III; structure of the complete IgA-molecule.";  
RL Hoppe-Seyler's Z. Physiol. Chem. 356:1337-1342(1975).  
CC -I- MISCELLANEOUS: THE SEQUENCE OF THE C REGION IS ALSO GIVEN.  
DR PIR; A02045; ALHUTR.  
DR InterPro; IPR003006; Ig\_MHC.  
DR InterPro; IPR003596; Ig\_v.  
DR Pfam; PF00047; ig; 1.  
DR SMART; SM00406; IGV; 1.  
KW Immunoglobulin V region.  
FT MOD\_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.  
FT NON\_TER 122 122  
SQ SEQUENCE 122 AA; 13472 MW; 2E21A11DA04D80F9 CRC64;

alignment\_scores:

Quality: 341.00 Length: 129  
Ratio: 3.410 Gaps: 3  
Percent Similarity: 77.519 Percent Identity: 52.713

alignment\_block:

US-09-742-690-1 x HV3A\_HUMAN ..

Align seg 1/1 to: HV3A\_HUMAN from: 1 to: 122

```
12 CAGGTGCAGCTGCAGGAGTCAAGGGGAGGATTGGTGCAGGCTGGGGGCTC 61
|||||
1 GluValGlnLeuValGlnSerGlyGlyGlyLeuValLysProGlyGlyse 17
62 TCTGAGACTCTCTGTGCAGCCTCGGGACGCGCCACACAGTGGTCATGGTC 111
|||||
17 rLeuArgLeuSerCysValAlaSerGly.....PheSerPheArg 31
112 ACTATGGTATGGGCTGTCCGCCAGGTTCCAGGGAAGGAGCGTGAGTTT 161
:::|||||
31 spPheTyrMetSerTrpIleArg***ThrProGlyLysGlyLeu***Trp 47
```

```
162 GTCGCAGCTATTAGGTGGAGTGTAAAGAGACATGGTATAAAGACTCCGT 211
||||| ||| ||||||| :||| |||||||
48 ValSerTyrIleGlyGlySerGlySerThrLeuTyrTyrAlaAspSerVa 64
||||| ||| ||||||| :||| |||||||
212 GAAGGGCCGATTACCATCTCCAGAGATAACGCCAAGACTACGGTTATC 261
||||| ||||||| ||||||| ||||||| ||||||| |||||||
64 lLysGlyArgPheThrIleSerArgAspAsnAlaGlnLysSerLeuTyrL 81
||||| ||||||| ||||||| ||||||| ||||||| |||||||
262 TCCAAATGAACAGCCCTGAACCTGAAGATACGGCCGCTTATTATTGTGCC 311
||||| ||||||| ||||||| ||||||| ||||||| |||||||
81 eu***Met***SerLeuArgThr*****ThrAlaValTyrTyrCysAla 97
||||| ||||||| ||||||| ||||||| ||||||| |||||||
312 GCTCGACCGGTCCGGTGGATGATATTTCCCTGCCGGTT...GGGTTTGA 358
||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
98 Ala.....Thr*****Phe***trpSerThrPheSerLeu** 110
||||| ||||||| ||||||| ||||||| ||||||| |||||||
359 CTAATGGGGCCAGGGGACCCAGGTCAACCGTCTCTCTCA 395
:||||| ||||||| ||||||| ||||||| ||||||| |||||||
110 *TyrTrpGly***Gly***LeuValThrValSerSer 122
||||| ||||||| ||||||| ||||||| ||||||| |||||||
```

seq\_name: SwissProt\_40:HV33\_MOUSE

```
seq_documentation_block:
ID HV33_MOUSE STANDARD; PRT; 115 AA.
AC P01802;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig heavy chain V-III region W3082.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE.
RX MEDLINE=82099361; PubMed=6798111;
RA Johnson N., Slankard J., Paul L., Hood L.;
RT "The complete V domain amino acid sequences of two myeloma inulin-
RT binding proteins."
RL J. Immunol. 128:302-307(1982).
CC -I- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM A MYELOMA PROTEIN THAT
CC BINDS INULIN.
DR PIR; D92811; AVMS82.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; IGv; 1.
KW Immunoglobulin V region.
FT DISULFID 22 98
FT NON_TER 115 115 BY SIMILARITY.
SQ SEQUENCE 115 AA; 12887 MW; 9B4517648C121C5A CRC64;
```

alignment\_scores:
Quality: 340.50 Length: 130
Ratio: 3.510 Gaps: 3
Percent Similarity: 74.615 Percent Identity: 54.615

alignment\_block:
US-09-742-690-1 x HV33\_MOUSE ..

Align seg 1/1 to: HV33\_MOUSE from: 1 to: 115

```
12 CAGGTGCAGCTGCAGGAGTCAGGGGAGGAGTGGTGCAGGCTGGGGGCTC 61
:||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
1 GluValLysLeuGluGluSerGlyGlyLeuValGlnProGlyGlySe 17

62 TCTGAGACTCTCTCTGTGCAGCCTCGGACGCGCCACCAGTGGTCTATGGTC 111
|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
17 rMetLysLeuSerCysValAlaSerGlyPheThrPheSer.....A 31

112 ACTATGGTATGGGCTGGTTCCGGCCAGGTTCACAGGAAGGAGCGTGAGTTT 161
```

```
:||| |||:|:| ||||| ||| ||| |||:|:|
31 snTyrTrpMetAsnTrpValArgGlnSerProGluLysGlyLeuGluTrp 47
||||| ||| ||||||| :||| |||||||
162 GTCGCAGCTATTAGGTGGAGTGTAAAGAG.....ACATGGTATAAAGA 205
||||| ||||| :|||:|:| ||| ||| :|
48 ValAlaGluIleArgLeuLysSerHisAsnTyrAlaThrHisTyrAlaGl 64
||||| ||||||| ||||||| ||||||| ||||||| |||||||
206 CTCCGTGAAGGGCCGATTACCATCTCCAGAGATAACGCCAAGACTACGG 255
:||||| ||||||| ||||||| ||||||| ||||||| |||||||
64 uSerValLysGlyArgPheThrIleSerArgAspSerLysSerSerV 81
||||| ||||||| ||||||| ||||||| ||||||| |||||||
256 TTTATCTGCAAATGAACAGCCTGAAACCTGAAGATACGGCCGCTTATTAT 305
||||| ||||||| ||||||| ||||||| ||||||| |||||||
81 alTyrLeuArgMetAsnAsnLeuArgProGluAspThrGlyIleTyrTyr 97
||||| ||||||| ||||||| ||||||| ||||||| |||||||
306 TGTGCCGCTCGACCGGTCCGGTGGATGATATTTCCCTGCCGGTGGGTT 355
|||:|:| :||| :||| :||| :||| :||| :||| :||| :|||
98 CysThr.....ThrGlyPh 102

356 TGACTACTGGGGCCAGGGACCCAGGTCAACCGTCTCTCTCA 395
| ||||| ||||||| ||||||| ||||||| ||||||| |||||||
102 eAlaTyrTrpGlyGlnGlyThrLeuValThrValSerAla 115
```

seq\_name: SwissProt\_40:HV3L\_HUMAN

```
seq_documentation_block:
ID HV3L_HUMAN STANDARD; PRT; 119 AA.
AC P01773;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig heavy chain V-III region BUR.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE (MYELOMA PROTEIN BUR).
RX MEDLINE=79151016; PubMed=107164;
RA Putnam F.W., Liu Y.-S.V., Low T.L.K.;
RT "Primary structure of a human IgA1 immunoglobulin. IV. Streptococcal
RT IgA1 protease, digestion, Fab and Fc fragments, and the complete
RT amino acid sequence of the alpha 1 heavy chain."
RL J. Biol. Chem. 254:2865-2874(1979).
DR PIR; A02056; ALHUBR.
DR HSSP; P01772; 2IG2.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; IGv; 1.
KW Immunoglobulin V region; Glycoprotein.
FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
FT DISULFID 22 96
FT CARBOHYD 28 28 N-LINKED (GLCNAC. . .).
FT NON_TER 119 119
SQ SEQUENCE 119 AA; 12981 MW; 12A709A75344D024 CRC64;
```

alignment\_scores:
Quality: 340.50 Length: 127
Ratio: 3.584 Gaps: 2
Percent Similarity: 74.803 Percent Identity: 55.906

alignment\_block:
US-09-742-690-1 x HV3L\_HUMAN ..

Align seg 1/1 to: HV3L\_HUMAN from: 1 to: 119

```
12 CAGGTGCAGCTGCAGGAGTCAGGGGAGGAGTGGTGCAGGCTGGGGGCTC 61
||||| ||||| ||||||| ||||||| ||||||| ||||||| ||
1 GluValGlnLeuValGluSerGlyGlyValGlnAlaGlyThrSe 17

62 TCTGAGACTCTCTCTGTGCAGCCTCGGACGCGCCACCAGTGGTCTATGGTC 111
```

```
|||||
17 rLeuArgLeuSerCysThrAlaSerAlaPheAsnLeuSer.....A 31
112 ACTATGGTATGGGCTGGTCCCGCAGGTTCCAGGGAAGGAGCGTGAGTTT 161
   |||:||||| ||| |||:||||| |||:||||| |||:|||||
31 sPTyRAlaMetHisTrpValArgGlnAlaProGlyLysGlyLeu**Trp 47
162 GTCGCAGCTATTAGTGGAGTGGTAAAGAGACATGCTATAAAGACTCCGT 211
   ||| ||| :|||:||||| |||:||||| |||:|||||
48 ValAlaLeuIleSerTyrglyGlySer**ThrTyTyTyAlaAspSerVa 64
212 GAAGGGCCGATTCCACCATCTCCAGAGATAACGCCAAGACTACGGTTTATC 261
   |||:||||| |||:||||| |||:||||| |||:|||||
64 lArgGlyArgPheThrIleSerArg**IleSerLys**ThrLeuTyRl 81
262 TGCAAATGAACAGCCTGAACCTGAAGATACGGCCGTTTATTATTGGCC 311
   |||:||||| |||:||||| |||:||||| |||:|||||
81 eu**MetLysThrLeuArgThrGluAspThrAlaValTyTyTyCysAla 97
312 GCTCGACCGGTCGGCGTGGATGATATTTCCCTGCCGCTGGGTTTGACTA 361
   |||:||||| |||:|||||
98 LysLeu.....IleAlaValAlaGlyThrArg**ph 108
362 CTGGGGCCAGGGGACCCAGGTCAACCGTCTCC 392
   :||||| ||| |||:|||||
108 eTrpGlyGlnGlyThrLeuValThrValSer 118
```

seq\_name: SwissProt\_40:HV32\_MOUSE

```
seq_documentation_block:
ID HV32_MOUSE STANDARD; PRT; 115 AA.
AC P01801;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig heavy chain V-III region J606.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE.
RX MEDLINE=82099361; PubMed=6798111;
RA Johnson N., Slankard J., Paul L., Hood L.;
RT "The complete V domain amino acid sequences of two myeloma inulin-
   binding proteins."
RL J. Immunol. 128:302-307(1982).
CC -!- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM A MYELOMA PROTEIN THAT
CC BINDS INULIN.
DR PIR; C92811; AVMS06.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; IGv; 1.
KW Immunoglobulin V region.
FT DISULFID 22 98 BY SIMILARITY.
FT NON_TER 115 115
SQ SEQUENCE 115 AA; 12810 MW; B67AD6638A121A5F CRC64;
```

alignment\_scores:  
Quality: 337.50 Length: 130  
Ratio: 3.479 Gaps: 3  
Percent Similarity: 74.615 Percent Identity: 54.615

alignment\_block:

US-09-742-690-1 x HV32\_MOUSE ..

Align seg 1/1 to: HV32\_MOUSE from: 1 to: 115

```
12 CAGGTGCAGCTGCAGGAGTCAGGGGAGGATTGGTGCAGCGCTGGGGGCTC 61
   :|||:||||| |||:||||| |||:||||| |||:|||||
1 GluValLysLeuGluGluSerGlyGlyLeuValGlnProGlyGlyse 17
```

```
62 TCTGAGACTCTCCTGCGACGCTCGGGACGCGCCACCAGTGGTCAATGGTC 111
   |||:||||| |||:||||| |||:||||| |||:|||||
17 rMetLysLeuSerCysValAlaSerGlyPheThrPheSer.....A 31
112 ACTATGGTATGGGCTGGTCCCGCAGGTTCCAGGGAAGGAGCGTGAGTTT 161
   :||| |||:||||| |||:||||| |||:|||||
31 snTyRTrpMetAsnTrpValArgGlnSerProGluLysGlyLeuGluTrp 47
162 GTCGCAGCTATTAGTGGAGTGGTAAAGAG.....ACATGGTATAAAGA 205
   ||| ||| :|||:||||| |||:||||| |||:|||||
48 ValAlaGluIleArgLeuLysSerAsnAsnTyRAlaThrHisTyRAlaGl 64
206 CTCCGTGAAGGGCCGATTCCACCATCTCCAGAGATAACGCCAAGACTACGG 255
   :||||| |||:||||| |||:||||| |||:|||||
64 uSerValLysGlyArgPheThrIleSerArgAspSerLysSerServ 81
256 TTTATCTGCAAAATGAACAGCCTGAAACCTGAAGATACGGCCGTTTATTAT 305
   |||:||||| |||:||||| |||:||||| |||:|||||
81 alTyRLeuGlnMetAsnAsnLeuArgAlaGluAspThrGlyIleTyTyR 97
306 TGTGCCGCTCGACCGGTCGGCGTGGATGATATTTCCCTGCCGCTGGGTT 355
   |||:||||| |||:|||||
98 CysThr.....ThrGlyPh 102
356 TGACTACTGGGGCCAGGGGACCCAGGTCAACCGTCTCCTCA 395
   | ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
102 eAlaTyRTrpGlyGlnGlyThrLeuValThrValSerAla 115
```







Quality: 407.00 Length: 128  
Ratio: 3.913 Gaps: 2  
Percent Similarity: 81.250 Percent Identity: 64.844

alignment\_block:

US-09-742-690-1 x S20782 ..

Align seg 1/1 to: S20782 from: 1 to: 124

```
12 CAGGTGCAGTGCAGGAGTCAGGGGGGAGGATGGTGCAGGCTGGGGGCTC 61
:::||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1 GluValGlnLeuLeuGluSerGlyGlyGlyLeuValGlnProGlyGlySe 17
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
62 TCTGAGACTCTCCTGTGCAGCCTCGGGACGCGCCACCAGTGGTTCATGGTC 111
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
17 rLeuArgLeuSerCysAlaAlaSerGlyPheThrPheSerSer..... 31
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
112 ACTATGGTATGGGCTGGTTCGCCAGGTTCAGGGAAGGAGCGTGAGTTT 161
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
32 ..TyrAlaMetAsnTrpValArgGlnAlaProGlyGluGlyLeuGluTrp 47
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
162 GTGCGCAGCTATTAGTGGAGTGGTAAAGACAGATGGTATAAAGACTCCGT 211
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
48 ValSerThrIleSerGlySerGlyAspSerThrTyrTyrAlaAspSerVa 64
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
212 GAAGGGCCGATTACCATCTCCAGAGATAACGCCAAGACTACGGTTTATC 261
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
64 lLysGlyArgPheThrIleSerArgAspAsnSerLysAsnThrLeuTyrL 81
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
262 TGCAAAATGAACAGCCTGAAACCTGAAGATACGGCGGTTTATTATTGTC 311
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
81 euGlnMetAsnSerLeuArgAlaGluAspThrAlaValTyrTyrCysAla 97
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
312 GCTCGACCGTCCGGTGGATGATATTTCCCTGCGGGTGGGTTTGACTA 361
::: ::: ::: ::: ::: ||||| ||||| ||||| |||||
98 LysGluArgIleAlaIlePheGlyValValIlePro...HisPheAspTy 113
::: ::: ::: ::: ::: ||||| ||||| ||||| |||||
362 CTGGGGCCAGGGACCCAGGTCAACCGTCTCCRCA 395
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
113 rTrpGlyGlnGlyThrLeuValThrValSerSer 124
```

seq\_name: pir2:S05271

seq\_documentation\_block:

Ig heavy chain precursor - human (fragment)  
C;Species: Homo sapiens (man)  
C;Date: 30-Jun-1992 #sequence\_revision 30-Jun-1992 #text\_change 16-Aug-1996  
C;Accession: S05271; S04602  
R;Kishimoto, T.  
submitted to the EMBL Data Library, March 1989  
A;Reference number: S05270  
A;Accession: S05271  
A;Molecule type: mRNA  
A;Residues: 1-160 <KIS1>  
A;Cross-references: EMBL:X14584  
R;Kishimoto, T.; Okajima, H.; Okumoto, T.; Taniguchi, M.  
Nucleic Acids Res. 17, 4385, 1989  
A;Title: Nucleotide sequences of the cDNAs encoding the V-regions of H- and L-chains of  
A;Reference number: S04601; MUID:89296497  
A;Accession: S04602  
A;Molecule type: mRNA  
A;Residues: 1-144 <KIS2>  
A;Cross-references: EMBL:X14584  
C;Superfamily: immunoglobulin V region; immunoglobulin homology  
C;Keywords: heterotetramer; immunoglobulin  
F;1-19/Domain: signal sequence #status predicted <SIG>  
F;20-160/Product: Ig heavy chain (fragment) #status predicted <MAT>  
F;34-117/Domain: immunoglobulin homology <IMM>

alignment\_scores:

Quality: 406.50 Length: 130  
Ratio: 4.025 Gaps: 1

Percent Similarity: 77.692 Percent Identity: 66.923

alignment\_block:

US-09-742-690-1 x S05271 ..

Align seg 1/1 to: S05271 from: 1 to: 160

```
12 CAGGTGCAGTGCAGGAGTCAGGGGGGAGGATGGTGCAGGCTGGGGGCTC 61
:::||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
20 GluValGlnLeuLeuGluSerGlyGlyGlyLeuValGlnProGlyGlySe 36
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
62 TCTGAGACTCTCCTGTGCAGCCTCGGGACGCGCCACCAGTGGTTCATGGTC 111
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
36 rLeuArgLeuSerCysAlaAlaSerGlyPheThrPheSer.....T 50
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
112 ACTATGGTATGGGCTGGTTCGCCAGGTTCAGGGAAGGAGCGTGAGTTT 161
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
50 hrTyrAlaMetSerTrpValArgGlnAlaProGlyLysGlyLeuGluTrp 66
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
162 GTCGACAGCTATTAGTGGAGTGGTAAAGACAGACATGGTATAAAGACTCCGT 211
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
67 ValSerAlaIleSerGlySerGlyGlySerThrTyrTyrAlaAspSerVa 83
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
212 GAAGGGCCGATTACCATCTCCAGAGATAACGCCAAGACTACGGTTTATC 261
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
83 lLysGlyArgPheThrIleSerArgAspAsnSerLysAsnThrLeuTyrL 100
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
262 TGCAAAATGAACAGCCTGAAACCTGAAGATAACGGCGGTTTATTATTGTC 311
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
100 euGlnMetAsnSerLeuArgAlaGluAspThrAlaValTyrTyrCysAla 116
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
312 GCTCGACCGTCCGGTGGATGATATTTCCCTGCGGGTGGGTTTGACTA 361
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
117 LysAlaValValArgGlyValIleSerTyrTyrTyrGlyMetAspVa 133
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
362 CTGGGGCCAGGGACCCAGGTCAACCGTCTCTCAGGATCT 401
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
133 lTrpGlyGlnGlyThrThrValThrValSerSerGlySer 146
```

seq\_name: pir2:A60943

seq\_documentation\_block:

Ig heavy chain precursor V region (clone HN.14) - human  
C;Species: Homo sapiens (man)  
C;Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 17-Mar-1999  
C;Accession: A60943; A48165  
R;Desai, R.; Spatz, L.; Matsuda, T.; Ilyas, A.A.; Berman, J.E.; Alt, F.W.; Kabat, E.A  
J. Neuroimmunol. 30, 245, 1990  
A;Title: Molecular cloning of a human immunoglobulin heavy chain variable (V-H) regio  
A;Reference number: A60943; MUID:91036050  
A;Accession: A60943  
A;Status: not compared with conceptual translation  
A;Molecule type: mRNA  
A;Residues: 1-151 <DES>  
R;Desai, R.; Spatz, L.; Matsuda, T.; Ilyas, A.A.; Berman, J.E.; Alt, F.W.; Kabat, E.A  
J. Neuroimmunol. 26, 35-41, 1990  
A;Title: Molecular cloning of a human immunoglobulin heavy chain variable (V-H) regio  
A;Reference number: A48165; MUID:90094677  
A;Accession: A48165  
A;Molecule type: mRNA  
A;Residues: 1-36,'M',38-62,'AR',67-151 <DE2>  
A;Note: this sequence has been corrected in reference A60943  
C;Superfamily: immunoglobulin V region; immunoglobulin homology  
C;Keywords: heterotetramer; immunoglobulin  
F;34-117/Domain: immunoglobulin homology <IMM>

alignment\_scores:

Quality: 406.00 Length: 130  
Ratio: 3.904 Gaps: 2  
Percent Similarity: 80.000 Percent Identity: 65.385

alignment\_block:



212 GAAGGGCCGATTCCACCATCTCCAGAGATAACGCCAAGACTACGGTTTATC 261  
|||||  
83 lLysGlyArgPheThrIleSerArgAspAsnSerLysAsnThrLeuTyrL 100  
262 TGCAAATGAACAGCCTGAAACCTGAAGATACGCCCGTTTATTATTGTGCC 311  
|||||  
100 euGlnMetAsnSerLeuArgAlaGluAspThrAlaValTyrTyrCysAla 116  
312 GCTCGACCGGTCCGGTGGATGATATTTCCCTGCCGGTTGGGTTTGACTA 361  
|||  
117 LysCysPropheAlaGlyGlySerProSer.....PheAspTy 129  
362 CTGGGGCCAGGGGACCCAGGTACCCGTCCTCTCTCA 395  
|||||  
129 rTrpGlyGlnGlyThrLeuValThrValSerSer 140  
seq\_name: pir2:S31108  
seq\_documentation\_block:  
Ig heavy chain - human  
C;Species: Homo sapiens (man)  
C;Date: 02-Dec-1993 #sequence\_revision 26-May-1995 #text\_change 17-Mar-1999  
C;Accession: S31108  
R;Raaphorst, F.M.; Timmers, E.; Kenter, M.J.H.; van Tol, M.J.D.; Vossen, J.M.; Schuurman  
Eur. J. Immunol. 22, 247-251, 1992  
A;Title: Restricted utilization of germ-line V(H)3 genes and short diverse third complement  
A;Reference number: S31104; MUID:92111633  
A;Accession: S31108  
A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
A;Molecule type: mRNA  
A;Residues: 1-119 <RAA>  
A;Cross-references: EMBL:X62956  
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, October 1991  
C;Superfamily: immunoglobulin V region; immunoglobulin homology  
C;Keywords: heterotetramer; immunoglobulin  
F;15-98/Domain: immunoglobulin homology <IMM>

alignment\_scores:  
Quality: 404.50 Length: 128  
Ratio: 4.045 Gaps: 2  
Percent Similarity: 78.125 Percent Identity: 67.188

alignment\_block:  
US-09-742-690-1 x S31108 ..  
Align seg 1/1 to: S31108 from: 1 to: 119

12 CAGGTGCAGCTGCAGAGTCAGGGGAGGATTGGTCAGGCTGGGGGCTC 61  
:::|||||  
1 GluValGlnLeuLeuGluSerGlyGlyGlyLeuValGlnProGlyGlySe 17  
62 TCTGAGACTCTCCTGTGCAGCCTCGGGACGCCACCAGTGGTCATGGTC 111  
|||||  
17 rLeuArgLeuSerCysAlaAlaSerGlyPheThrPheSerSer..... 31  
112 ACTATGGTATGGCTGGTTCGCCAGGTTCCAGGGAAGGAGCGTGAGTTT 161  
|||:::|||||  
32 ..TyrAlaMetSerTrpValArgGlnAlaProGlyLysGlyLeuGluTrp 47  
162 GTCGCAGCTATTAGTGGAGTGGTAAAGAGACATGGTATAAAGACTCCGT 211  
|||:::|||||  
48 ValSerAlaIleSerGlySerGlyGlySerThrTyrTyrAlaAspSerVa 64  
212 GAAGGGCCGATTCCACCATCTCCAGAGATAACGCCAAGACTACGGTTTATC 261  
|||||  
64 lLysGlyArgPheThrIleSerArgAspAsnSerLysAsnThrLeuTyrL 81  
262 TGCAAATGAACAGCCTGAAACCTGAAGATACGCCCGTTTATTATTGTGCC 311  
|||||  
81 euGlnMetAsnSerLeuArgAlaGluAspThrAlaValTyrTyrCysAla 97  
312 GCTCGACCGGTCCGGTGGATGATATTTCCCTGCCGGTTGGGTTTGACTA 361

98 .....LysAspArgArgLeuThrGlyThrPheAspTy 108  
|||  
362 CTGGGGCCAGGGGACCCAGGTACCCGTCCTCTCTCA 395  
|||||  
108 rTrpGlyGlnGlyThrLeuValThrValSerSer 119  
seq\_name: pir2:S31107  
seq\_documentation\_block:  
Ig heavy chain - human  
C;Species: Homo sapiens (man)  
C;Date: 02-Dec-1993 #sequence\_revision 26-May-1995 #text\_change 17-Mar-1999  
C;Accession: S31107  
R;Raaphorst, F.M.; Timmers, E.; Kenter, M.J.H.; van Tol, M.J.D.; Vossen, J.M.; Schuur  
Eur. J. Immunol. 22, 247-251, 1992  
A;Title: Restricted utilization of germ-line V(H)3 genes and short diverse third comp  
A;Reference number: S31104; MUID:92111633  
A;Accession: S31107  
A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
A;Molecule type: mRNA  
A;Residues: 1-119 <RAA>  
A;Cross-references: EMBL:X62955  
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, October 1991  
C;Superfamily: immunoglobulin V region; immunoglobulin homology  
C;Keywords: heterotetramer; immunoglobulin  
F;15-98/Domain: immunoglobulin homology <IMM>

alignment\_scores:  
Quality: 403.50 Length: 128  
Ratio: 4.035 Gaps: 2  
Percent Similarity: 78.125 Percent Identity: 66.406

alignment\_block:  
US-09-742-690-1 x S31107 ..  
Align seg 1/1 to: S31107 from: 1 to: 119

12 CAGGTGCAGCTGCAGAGTCAGGGGAGGATTGGTCAGGCTGGGGGCTC 61  
:::|||||  
1 GluValGlnLeuValGluSerGlyGlyGlyLeuValGlnProGlyGlySe 17  
62 TCTGAGACTCTCCTGTGCAGCCTCGGGACGCCACCAGTGGTCATGGTC 111  
|||||  
17 rLeuArgLeuSerCysAlaAlaSerGlyPheThrPheSerSer..... 31  
112 ACTATGGTATGGCTGGTTCGCCAGGTTCCAGGGAAGGAGCGTGAGTTT 161  
|||:::|||||  
32 ..TyrAlaMetSerTrpValArgGlnAlaProGlyLysGlyLeuGluTrp 47  
162 GTCGCAGCTATTAGTGGAGTGGTAAAGAGACATGGTATAAAGACTCCGT 211  
|||:::|||||  
48 ValSerAlaIleSerGlySerGlyGlySerThrTyrTyrAlaAspSerVa 64  
212 GAAGGGCCGATTCCACCATCTCCAGAGATAACGCCAAGACTACGGTTTATC 261  
|||||  
64 lLysGlyArgPheThrIleSerArgAspAsnSerLysAsnThrLeuTyrL 81  
262 TGCAAATGAACAGCCTGAAACCTGAAGATACGCCCGTTTATTATTGTGCC 311  
|||||  
81 euGlnMetAsnSerLeuArgAlaGluAspThrAlaValTyrTyrCysAla 97  
312 GCTCGACCGGTCCGGTGGATGATATTTCCCTGCCGGTTGGGTTTGACTA 361  
|||  
98 LysAspPro.....GlyAlaSerTyrTyrPheAspTy 108  
362 CTGGGGCCAGGGGACCCAGGTACCCGTCCTCTCTCA 395  
|||||  
108 rTrpGlyGlnGlyThrLeuValThrValSerSer 119

seq\_name: pir2:S31588







```
||||| ||| |||:||||| |||:
32 ..TyrGlyMetHisTrpValLeuGlnAlaProGlyLysGlyLeuGluTrp 47
|||||:||||| :|||:|||||:|||||
162 GTCGAGCTATTAGGTGGAGTGCTAAAGAGACATGGTATAAAGACTCCGT 211
|||||:||||| :|||:|||||:|||||
48 ValAlaValIleTrpTyrAspGlySerAsnLysTyrTyrAlaAspSerVa 64
|||||:|||||:|||||:|||||:|||||
212 GAAGGCGCGATTCAACCATCTCCAGAGATAACGCCAAGACTACGGTTTATC 261
|||||:|||||:|||||:|||||:|||||
64 lLysGlyArgPheThrIleSerArgAspAsnSerLysAsnThrLeuTyrL 81
|||||:|||||:|||||:|||||:|||||
262 TGCAATGAACAGCCTGAAACCTGAAGATACGGCCGTTTATATTGTGCC 311
|||||:|||||:|||||:|||||:|||||
81 euGlnMetAsnSerLeuArgAlaGluAspThrAlaValTyrTyrCys... 96
|||||:|||||:|||||:|||||:|||||
312 GCTCGACCGGTCCCGGTGGATGATATTCCCTGCCGGTTGGGTTTGACTA 361
|||||:|||||:|||||:|||||:|||||
97 AlaArgAspArgArgLeu...ThrIleAlaAlaAlaGlyAsnPheAspTy 112
|||||:|||||:|||||:|||||:|||||
362 CTGGGGCCAGGGGACCCAGGTACCGTCTCCTCAGGATCT 401
|||||:|||||:|||||:|||||:|||||
112 rTrpGlyGlnGlyThrLeuAlaThrValSerSerGlySer 125
```

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 13, 2002, 07:00:02 ; Search time 77.58 Seconds  
(without alignments)  
1766.737 Million cell updates/sec

Title: US-09-742-690-1  
Perfect score: 558  
Sequence: 1 tcgagaaaagacaggtgcag.....ctcccagtgctgtaataag 558

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_NA.\*  
1: /cgn2\_6/ptodata/2/ina/5A\_COMB.seq.\*  
2: /cgn2\_6/ptodata/2/ina/5B\_COMB.seq.\*  
3: /cgn2\_6/ptodata/2/ina/6A\_COMB.seq.\*  
4: /cgn2\_6/ptodata/2/ina/6B\_COMB.seq.\*  
5: /cgn2\_6/ptodata/2/ina/PCTUS\_COMB.seq.\*  
6: /cgn2\_6/ptodata/2/ina/backfiles1.seq.\*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	198	35.5	867	4 US-08-862-124-16	Sequence 16, Appl
2	196.6	35.2	375	4 US-09-240-274-93	Sequence 93, Appl
3	196.4	35.2	867	4 US-08-862-124-18	Sequence 18, Appl
4	196	35.1	918	4 US-08-862-124-13	Sequence 13, Appl
5	196	35.1	918	4 US-08-862-124-15	Sequence 15, Appl
6	195	34.9	375	4 US-09-240-274-78	Sequence 78, Appl
7	191	34.2	378	4 US-09-240-274-192	Sequence 192, App
8	185.6	33.3	903	4 US-09-273-839A-12	Sequence 12, Appl
9	185.6	33.3	913	4 US-09-273-839A-10	Sequence 10, Appl
10	185.4	33.2	381	4 US-09-240-274-88	Sequence 88, Appl
11	183	32.8	360	2 US-08-428-197-21	Sequence 21, Appl
12	183	32.8	360	5 PCT-US93-10555-21	Sequence 21, Appl
13	182.2	32.7	372	4 US-09-240-274-71	Sequence 71, Appl
14	182.2	32.7	372	4 US-09-240-274-76	Sequence 76, Appl
15	182	32.6	378	4 US-09-240-274-79	Sequence 79, Appl
16	181.4	32.5	360	2 US-08-428-197-19	Sequence 19, Appl
17	181.4	32.5	360	5 PCT-US93-10555-19	Sequence 19, Appl
18	180.8	32.4	384	4 US-09-240-274-70	Sequence 70, Appl
19	180.6	32.4	372	4 US-09-240-274-73	Sequence 73, Appl
20	180.6	32.4	372	4 US-09-240-274-74	Sequence 74, Appl
21	180.6	32.4	372	4 US-09-240-274-75	Sequence 75, Appl
22	180.6	32.4	375	4 US-09-240-274-89	Sequence 89, Appl
23	180.4	32.3	378	4 US-09-240-274-193	Sequence 193, App
24	179.8	32.2	360	2 US-08-428-197-23	Sequence 23, Appl
25	179.8	32.2	360	2 US-08-428-197-25	Sequence 25, Appl
26	179.8	32.2	360	2 US-08-428-197-27	Sequence 27, Appl
27	179.8	32.2	360	5 PCT-US93-10555-23	Sequence 23, Appl

28	179.8	32.2	360	5 PCT-US93-10555-25	Sequence 25, Appl
29	179.8	32.2	360	5 PCT-US93-10555-27	Sequence 27, Appl
30	179.4	32.2	406	1 US-08-471-780C-99	Sequence 99, Appl
31	179.4	32.2	406	1 US-08-467-282B-99	Sequence 99, Appl
32	179.4	32.2	406	2 US-08-471-282A-99	Sequence 99, Appl
33	179.4	32.2	406	2 US-08-466-710C-99	Sequence 99, Appl
34	179.4	32.2	406	3 US-08-471-284B-99	Sequence 99, Appl
35	179.4	32.2	406	3 US-08-468-739C-99	Sequence 99, Appl
36	179.2	32.1	384	4 US-09-240-274-185	Sequence 185, App
37	179	32.1	375	4 US-09-240-274-90	Sequence 90, Appl
38	179	32.1	375	4 US-09-240-274-183	Sequence 183, Appl
39	179	32.1	381	4 US-09-240-274-87	Sequence 87, Appl
40	178.8	32.0	378	4 US-09-240-274-187	Sequence 187, App
41	178.8	32.0	514	3 US-08-545-809A-23	Sequence 23, Appl
42	178.6	32.0	369	2 US-08-428-197-37	Sequence 37, Appl
43	178.6	32.0	369	5 PCT-US93-10555-37	Sequence 37, Appl
44	178	31.9	375	4 US-09-240-274-91	Sequence 91, Appl
45	177.8	31.9	432	1 US-08-026-320A-1	Sequence 1, Appl

ALIGNMENTS

RESULT 1  
US-08-862-124-16  
; Sequence 16, Application US/08862124  
; Patent No. 6207153  
; GENERAL INFORMATION:  
; APPLICANT: Dan, Michael D.  
; APPLICANT: Maiti, Pradip K.  
; APPLICANT: Kaplan, Howard A.  
; TITLE OF INVENTION: ANTIGEN BINDING FRAGMENTS H11, THAT  
; TITLE OF INVENTION: SPECIFICALLY DETECT CANCER CELLS, NUCLEOTIDES ENCODING THE  
; TITLE OF INVENTION: FRAGMENTS, AND USE THEREOF FOR THE PROPHYLAXIS AND  
; TITLE OF INVENTION: DETECTION OF CANCERS  
; NUMBER OF SEQUENCES: 28  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Morrison & Foerster LLP  
; STREET: 755 Page Mill Road  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94304-1018  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/862,124  
; FILING DATE: 22-MAY-1997  
; CLASSIFICATION: 424  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Lehnhardt, Susan K.  
; REGISTRATION NUMBER: 33,943  
; REFERENCE/DOCKET NUMBER: 31608-20001.20  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (650) 813-5600  
; TELEFAX: (650) 494-0792  
; TELEX: 706141  
; INFORMATION FOR SEQ ID NO: 16:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 867 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: join(1..855, 862..867)  
US-08-862-124-16

Query Match 35.5%; Score 198; DB 4; Length 867;



Best Local Similarity 70.4%; Pred. No. 1.8e-54;  
Matches 283; Conservative 0; Mismatches 110; Indels 9; Gaps 1;

```

QY 10 gacaggtgcagctgcaggagtcaggggagggattggtcagggctggggctctctgagac 69
    |||||
Db 425 GACAGGTGCAGCTGGTGGAGTCTGGGGAGGCGTGGTCCAGCCTGGGAGGTCCCTGAGAC 484
    |||||

QY 70 tctcctgtgcagcctcgagcgcgccaccagtggtcatggtcactatggtatgggctggt 129
    |||||
Db 485 TCTCCTGTGCAGCCTCTGGATTCCCTTCAGAAG-----CTTGTCTATGCACCTGGG 535
    |||||

QY 130 tccgccaggttccaggaaggagcgtgagttgtcgcagctattagggtggagtggtaaag 189
    |||||
Db 536 TCCGCCAGGCTCTAGGCAAGGGGCTGGAGTGGGTGGCAGTTATATCATATGATGGAAGCA 595
    |||||

QY 190 agacatggtataaagactcccgtagggccgattccaccatctccagagataacgccaaaga 249
    |||||
Db 596 CTAATACTACGCAGACTCCGTGAAGGGCCGATTCACCATCTCCAGAGACACTTCCAAGA 655
    |||||

QY 250 ctacggtttatctgcaaatgaacagccctgaacacctgaacacctgaacacctgattattgtg 309
    |||||
Db 656 ACACGGTGTATCTAAATAATGAACAGCCTGAGAACTGAGACACGGCTGTCTATTACTGTG 715
    |||||

QY 310 ccgctcgaccggtccgggtggatgatatttccctgcgggttgggttggactactggggcc 369
    |||||
Db 716 CGAGAGATCAGAGCCTGTGGGTGACTATGACCACTACTACGGTTTGGACGCTCTGGGGCA 775
    |||||

QY 370 aggggacccaggtcacctcctcctcaggatctcatcaccatc 411
    |||||
Db 776 AAGGGACCACGGTCAACCGTCTCCTCAGGATCCGAACAAAAAC 817
    |||||

```

RESULT 2

US-09-240-274-93

; Sequence 93, Application US/09240274

; Patent No. 6255455

; GENERAL INFORMATION:

; APPLICANT: Siegel, Donald L.

; TITLE OF INVENTION: Rh(D)-BINDING PROTEINS AND MAGNETICALLY ACTIVATED CELL

; TITLE OF INVENTION: SORTING METHOD FOR PRODUCTION THEREOF

; FILE REFERENCE: 09596-42U2

; CURRENT APPLICATION NUMBER: US/09/240,274

; CURRENT FILING DATE: 1999-01-29

; EARLIER APPLICATION NUMBER: 60/081,380

; EARLIER FILING DATE: 1998-04-10

; EARLIER APPLICATION NUMBER: 60/028,550

; EARLIER FILING DATE: 1996-10-11

; NUMBER OF SEQ ID NOS: 224

; SOFTWARE: PatentIn ver. 2.0

; SEQ ID NO 93

; LENGTH: 375

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; OTHER INFORMATION: anti-Rh(D) chain D20

US-09-240-274-93

```

QY 193 catggtataaaagactccggtgaaggccgattccaccatctccagagataaacgccaagacta 252
Db 173 aggaatatgcagactccggtgaaggccgattccaccatctccagagataaacgccaagacta 232
QY 253 cgggtttatctgcaaatgaacagcctgaaacctgaagatacaggccgctttattattgtgccg 312
Db 233 cgctgtatctacaaatgaacagcctgagagccgagacacgctgtgtattactgtgcga 292
QY 313 ctgcagccggtccgctggtgatgatatctccctgccggttggtttgactactggtggccagg 372
Db 293 gagaagaagtgttcggggagttatcttattgtctcggaaagtttgactactggtggccagg 352
QY 373 ggacccaggtcacccgtctctca 395
Db 353 gaacctgttcacccgtctctca 375

RESULT 3
US-08-862-124-18/c
; Sequence 18, Application US/08862124
; Patent No. 6207153
; GENERAL INFORMATION:
; APPLICANT: Dan, Michael D.
; APPLICANT: Maiti, Pradip K.
; APPLICANT: Kaplan, Howard A.
; TITLE OF INVENTION: ANTIGEN BINDING FRAGMENTS H11, THAT
; TITLE OF INVENTION: SPECIFICALLY DETECT CANCER CELLS, NUCLEOTIDES EN
; TITLE OF INVENTION: FRAGMENTS, AND USE THEREOF FOR THE PROPHYLAXIS A
; TITLE OF INVENTION: DETECTION OF CANCERS
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morrison & Foerster LLP
; STREET: 755 Page Mill Road
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/862,124
; FILING DATE: 22-MAY-1997
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Lehnhardt, Susan K.
; REGISTRATION NUMBER: 33,943
; REFERENCE/DOCKET NUMBER: 31608-20001.20
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 813-5600
; TELEFAX: (650) 494-0792
; TELEX: 706141
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 867 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
;
US-08-862-124-18

```

||||| 383 TCTCTGTGCAGCCTCTGATTCCTTTCAGAAAG-----CTTTGCTATGCACTGGG 333  
QY 130 tccgccaggttccagggaagcgtgagttgttcgcagctattaggtggagtggttaaag 189  
Db 332 TCCGCCAGGCTCTAGCAAGGGCTGGAGTGGTGCGCAGTTATATCATATGATGGAAGCA 273  
QY 190 agacatggtataaagactccgtgaaggccgattcaccatctccagagataacgccaaga 249  
Db 272 CTAAATACTACGCAGACTCCGTGAAGGCCGATTCACCATCTCCAGAGACACTTCCAAGA 213  
QY 250 ctacggtttatctgcaaatgaacagcctgaacctgaagatacggcgtttattattgtg 309  
Db 212 ACACGGTGTATCTAAATAAGAACAGCCTGAGAACTGAGGACACGGCTGCTACTGTG 153  
QY 310 ccgctcgaccggtccgctggatgatatttccctgccggttggttgactactggggcc 369  
Db 152 CGAGAGATCAGAGCCTGTGGGTGACTATGACCACTACTACGGTTTGGTCTGCTGGGGCA 93  
QY 370 aggggaccaggtcacctctcctcaggatctcatcaccatc 411  
Db 92 AAGGGACCACGGTCAACCGTCTCCTCAGGATCCGAACAAAAAC 51  
RESULT 4  
US-08-862-124-13  
; Sequence 13, Application US/08862124  
; Patent No. 6207153  
; GENERAL INFORMATION:  
; APPLICANT: Dan, Michael D.  
; APPLICANT: Maiti, Pradip K.  
; APPLICANT: Kaplan, Howard A.  
; TITLE OF INVENTION: ANTIGEN BINDING FRAGMENTS H11, THAT  
; TITLE OF INVENTION: SPECIFICALLY DETECT CANCER CELLS, NUCLEOTIDES ENCODING THE  
; TITLE OF INVENTION: FRAGMENTS, AND USE THEREOF FOR THE PROPHYLAXIS AND  
; TITLE OF INVENTION: DETECTION OF CANCERS  
; NUMBER OF SEQUENCES: 28  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Morrison & Foerster LLP  
; STREET: 755 Page Mill Road  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94304-1018  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/862,124  
; FILING DATE: 22-MAY-1997  
; CLASSIFICATION: 424  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Lehnhardt, Susan K.  
; REGISTRATION NUMBER: 33,943  
; REFERENCE/DOCKET NUMBER: 31608-20001.20  
; TELEPHONE: (650) 813-5600  
; TELEFAX: (650) 494-0792  
; TELEX: 706141  
; INFORMATION FOR SEQ ID NO: 13:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 918 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: join(1..906, 913..918)  
US-08-862-124-13

Query Match 35.1%; Score 196; DB 4; Length 918;  
Best Local Similarity 70.2%; Pred. No. 8.2e-54;  
Matches 281; Conservative 0; Mismatches 110; Indels 9; Gaps 1;  
QY 12 caggtgcagctgcagagagtcagggggagagattggtgcaggtgggggctctctgagactc 71  
Db 478 CAGGTGCAGCTGGTGGAGTCTGGGGGAGGGCTGCCAGCCTGGGAGGTCCCTGAGACTC 537  
QY 72 tccgtgcagcctcgggacgcgccaccagtggtcatggtcactatggtatgggtggttc 131  
Db 538 TCCTGTGCAGCCTCTGGATTCCCTTTCAGAAG-----CTTTGCTATGCACTGGGTC 588  
QY 132 cgccaggttccagggaagcgtgagttgttcgagctattaggtggagtggttaaagag 191  
Db 589 CGCCAGGCTCTAGGCAAGGGCTGGAGTGGGTGGCAGTTTACCATCTCCAGAGACACTTCCAAGAAC 648  
QY 192 acatggtataaagactccgtgaaggccgattcaccatctccagagataacgccaagact 251  
Db 649 AAATACTACGCAGACTCCGTGAAGGCCGATTCACCATCTCCAGAGACACTTCCAAGAAC 708  
QY 252 acggtttatctgcaaatgaacagcctgaacctgaagatacggcgtttattattgtgcc 311  
Db 709 ACGGTGTATCTAAATAAGAACAGCCTGAGAACTGAGGACACGGCTGTCTATTACTGTGG 768  
QY 312 gctcgaccggtccgctggatgatatttccctgccggttggttgactactggggccag 371  
Db 769 AGAGATCAGAGCCTGTGGGTGACTATGACCACTACTACGGTTTGGACGCTCTGGGCAAA 828  
QY 372 gggaccaggtcacctctcctcaggatctcatcaccatc 411  
Db 829 GGGACCACGGTCAACCGTCTCCTCAGGATCCGAACAAAAAC 868  
RESULT 5  
US-08-862-124-15/c  
; Sequence 15, Application US/08862124  
; Patent No. 6207153  
; GENERAL INFORMATION:  
; APPLICANT: Dan, Michael D.  
; APPLICANT: Maiti, Pradip K.  
; APPLICANT: Kaplan, Howard A.  
; TITLE OF INVENTION: ANTIGEN BINDING FRAGMENTS H11, THAT  
; TITLE OF INVENTION: SPECIFICALLY DETECT CANCER CELLS, NUCLEOTIDES ENCODING THE  
; TITLE OF INVENTION: FRAGMENTS, AND USE THEREOF FOR THE PROPHYLAXIS AND  
; TITLE OF INVENTION: DETECTION OF CANCERS  
; NUMBER OF SEQUENCES: 28  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Morrison & Foerster LLP  
; STREET: 755 Page Mill Road  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94304-1018  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/862,124  
; FILING DATE: 22-MAY-1997  
; CLASSIFICATION: 424  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Lehnhardt, Susan K.  
; REGISTRATION NUMBER: 33,943  
; REFERENCE/DOCKET NUMBER: 31608-20001.20  
; TELEPHONE: (650) 813-5600  
; TELEFAX: (650) 494-0792  
; TELEX: 706141  
; INFORMATION FOR SEQ ID NO: 15:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 918 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: join(1..906, 913..918)  
US-08-862-124-15







```
Query Match 33.2%; Score 185.4; DB 4; Length 381;
Best Local Similarity 71.5%; Pred. No. 1.5e-50;
Matches 278; Conservative 0; Mismatches 96; Indels 15; Gaps 2;

QY 13 aggtgcagctgcaggagtcaggggaggttggtgcaggtggtgggctctctgagactct 72
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2 aggtgcagctgctcagtcggtggggaggtctgtgtacagcctgggggtccctcagactct 61

QY 73 cctgtgcagcctcgggacgcgccaccagtggtcatgtggtcactatggtatgggtgttcc 132
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 62 cctgtgcagcgtctggtgattcaccttcagtact-----tatggcatgactgggtcc 112

QY 133 gccaggttccagggaaggagcgtgagtttgcgcagctattaggtggagtggttaagaga 192
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 113 gccaggtccaggaagggtggtgagtggtggtgagttatatggtttgatggaagtaaga 172

QY 193 catggtataaagactccgtgaaggccgattcaccatctccagagataacgccaaagacta 252
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 173 gagactatgcagagtcctgtaaggccgattcaccatctccagagataacactccaaagaa 232

QY 253 cggtttatctgcaaatgaacagcctgaagccgattcaccatctccagagataacgtgtgccg 312
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 233 cactgtatctgcaaatgaacagcctgagagcggaggtcgtgtgtattactgtcgga 292

QY 313 ctgcagccgt-----ccgctggtgatgatattccctgcgggttggtttgactactggg 366
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 293 gagaaatgtgctgctggtggggggggcattcgatgatacaagtaactactttgactactggg 352

QY 367 gccagggaccaggtcacctgtctcctca 395
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 353 gccagggaaccctggtcacctgtctcctca 381

RESULT 11
US-08-428-197-21
; Sequence 21, Application US/08428197
; Patent No. 5891438
; GENERAL INFORMATION:
; APPLICANT: SILVERMAN, GREGG J.
; TITLE OF INVENTION: METHOD FOR STIMULATING PRODUCTION OF
; TITLE OF INVENTION: VARIABLE REGION GENE FAMILY RESTRICTED ANTIBODIES THROUGH
; TITLE OF INVENTION: VACCINATION WITH A B-CELL SUPERANTIGEN AND CONJUGATES
; TITLE OF INVENTION: THEREOF
; NUMBER OF SEQUENCES: 51
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Spensley Horn Jubas & Lubitz
; STREET: 1880 Century Park East - Suite 500
; CITY: Los Angeles
; STATE: California
; COUNTRY: USA
; ZIP: 90067
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/428,197
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/10555
; FILING DATE: 29-OCT-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Howells, Stacy L.
; REGISTRATION NUMBER: 34,842
; REFERENCE/DOCKET NUMBER: FD-2630
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 455-5100
; TELEFAX: (619) 455-5110
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
```

```
; LENGTH: 360 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; IMMEDIATE SOURCE:
; CLONE: SPA3-08
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..360
US-08-428-197-21

Query Match 32.8%; Score 183; DB 2; Length 360;
Best Local Similarity 72.2%; Pred. No. 8.4e-50;
Matches 275; Conservative 0; Mismatches 85; Indels 21; Gaps 2;

QY 12 caggtgcagctgcaggagtcaggggaggttggtgcaggtgggggctctctgagactc 71
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1 CAGGTGAAACTGCTCGAGTCTGGGGGAGGATGGGTACAGCCTGGGGGTCCCTGAGACTC 60

QY 72 tccgtgcagcctcgggacgcgccaccagtggtcatggtcactatggtatgggtgttcc 131
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 61 TCCTGTGCAGCCTCTGTGATTACCTTTAGCAGCCA-----TGCCATGAGCTGGGTC 111

QY 132 cgcaggttccagggaaggagcgtgagtttgcgcagctattaggtggagtggttaaagag 191
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 112 CGCCAGGCTCCAGGAAGGGCCTGGAGTGGGTCTCAGATATTAGTGCCAGTGGTGTAGC 171

QY 192 acatggtataaagactccgtgaaggccgattcaccatctccagagataacgccaaagact 251
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 172 ACATATTATGCAGACTCCGTGAAGGGCCGTTACCATCTCCAGAGACAATCCAAAGAAC 231

QY 252 acggtttatctgcaaatgaacagcctgaacacctgaagatacggcgtttattattgtgcc 311
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 232 ACGCTGTATTGCAATGAACAGCCTGAGAGCCGGAAGACACGGCCTTATATTACTGTGCG 291

QY 312 gctgcagccgttcgcgtggtgatgatatttccctgcgggttgggttgcactactggggccag 371
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 292 TCCAGCGCGCGGGGATGGGGGTACCTTCCC-----TTGACTACTGGGGGCCAG 339

QY 372 gggaccaggtcacctgtctcc 392
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 340 GGAACCTGGTCAACCGTCTCC 360

RESULT 12
PCT-US93-10555-21
; Sequence 21, Application PC/TUS9310555
; GENERAL INFORMATION:
; APPLICANT: SILVERMAN, GREGG J.
; TITLE OF INVENTION: METHOD FOR STIMULATING PRODUCTION OF
; TITLE OF INVENTION: VARIABLE REGION GENE FAMILY RESTRICTED ANTIBODIES THROUGH
; TITLE OF INVENTION: VACCINATION WITH A B-CELL SUPERANTIGEN AND CONJUGATES
; TITLE OF INVENTION: THEREOF
; NUMBER OF SEQUENCES: 51
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Spensley Horn Jubas & Lubitz
; STREET: 1880 Century Park East - Suite 500
; CITY: Los Angeles
; STATE: California
; COUNTRY: USA
; ZIP: 90067
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/10555
; FILING DATE: 29-OCT-1993
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
```

```
; NAME: Howells, Stacy L.
; REGISTRATION NUMBER: 34,842
; REFERENCE/DOCKET NUMBER: FD-2630
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 455-5100
; TELEFAX: (619) 455-5110
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 360 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; IMMEDIATE SOURCE:
; CLONE: Spa3-08
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..360
PCT-US93-10555-21

Query Match          32.8%; Score 183; DB 5; Length 360;
Best Local Similarity 72.2%; Pred. No. 8.4e-50;
Matches 275; Conservative 0; Mismatches 85; Indels 21; Gaps 2;

QY 12 caggtgcagctgcaggagtcagggggagagattggtgcaggctgggggtctctgagactc 71
      ||||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| |
Db 1 CAGGTGAACACTGCTCGAGTCTGGGGGAGGATTGGTACAGCCTGGGGGGTCCCTGAGACTC 60

QY 72 tcctgtgcagcctcgggacgcgccaccagtggtcatggtcactatggtatgggctggttc 131
      ||||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| |
Db 61 TCCTGTGCAGCCTCTGGATTACCTTTAGCAGCCA-----TGCCATGAGCTGGGTC 111

QY 132 cgccaggttccagggaaggagcgtgagtttctgcagctattaggtggagtggtaaagag 191
      ||||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| |
Db 112 CGCCAGGCTCCAGGGAAGGCCTGGAGTGGGTCTCAGATATTAGTGCCAGTGGTAGC 171

QY 192 acatggtataagactccgtgaaggccgattcaccatctccagagataacgccaaagact 251
      |||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| |
Db 172 ACATATTATGCAGACTCCGTGAAGGCCGCTCACCATCTCCAGAGACAAATCCAAAGAAC 231

QY 252 acggtttatctcaaatgaacagccctgaaacctgaagatacggccgtttattattgtgcc 311
      |||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| |
Db 232 ACGCTGATTTCGAAATGAACAGCCTGAGAGCCGAAGACACGGCCTTATATTACTGTGCG 291

QY 312 gctcgaccggtccgctggtgatgatatttccctgcgggttgggttgactactg99ggccag 371
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 292 TCCAGCGCGCGGGATGGGGGTACCTTCCC-----TTGACTACTGGGGCCAG 339

QY 372 gggaccaggtcacctgtctcc 392
      || |||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| |
Db 340 GGAACCTGGTCAACCGTCTCC 360

RESULT 13
US-09-240-274-71
; Sequence 71, Application US/09240274
; Patent No. 6255455
; GENERAL INFORMATION:
; APPLICANT: Siegel, Donald L.
; TITLE OF INVENTION: Rh(D)-BINDING PROTEINS AND MAGNETICALLY ACTIVATED CELL
; FILE REFERENCE: 09596-4202
; CURRENT APPLICATION NUMBER: US/09/240,274
; CURRENT FILING DATE: 1999-01-29
; EARLIER FILING DATE: 1998-04-10
; EARLIER APPLICATION NUMBER: 60/081,380
; EARLIER FILING DATE: 1996-10-11
; NUMBER OF SEQ ID NOS: 224
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 71
; LENGTH: 372

Query Match          32.8%; Score 182.2; DB 4; Length 372;
Best Local Similarity 71.3%; Pred. No. 1.6e-49;
Matches 273; Conservative 0; Mismatches 98; Indels 12; Gaps 2;

QY 13 aggtgcagctgcaggagtcagggggagattggtgcaggctggggctctctgagactct 72
      ||||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| |
Db 2 aggtgcagctgctcgagctctgggggagcgctggtccagcctgggaggtccctgagactct 61

QY 73 cctgtgcagcctcgggacgcgccaccagtggtcatggtcactatggtatgggctggttcc 132
      ||||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| |
Db 62 cctgtgcagcctcgtgattctccttcagtag-----ctatggcatgcaactgggtcc 112

QY 133 gccaggttccagggaaggagcgtgagtttctgcagctattaggtggagtggttaaagaga 192
      ||||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| |
Db 113 gccaggtccaggcaagggtggtgagtggtgagttatcatatcatatgatggacatcata 172

QY 193 catggtataaagactccgtgaaggccgattcaccatctccagagataacgccaaagacta 252
      | ||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| |
Db 173 aaaactatgcagactccgtgaaggccgattcaccatctccagagataatccaagaaa 232

QY 253 cggtttatctgcaaatgaacagccctgaaacctgaagatacggccgtttattattgtgccg 312
      || | || ||||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| |
Db 233 cgctgtacctgcaaatgaacagccctgagacctgaggaacacggctgtattactgtgcga 292

QY 313 ctcgaccggtccgcgtggtgatgatatttccctgcgggttgggttgactactg99ggccag 372
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 293 acctaaagggggaagtaactcgtcgtgc---gtctgttcccttgatatcttggggccag 349

QY 373 ggaccaggtcacctgtctctcca 395
      |||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| |
Db 350 ggacaatggtcacgtctcttcca 372
```

```
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: anti-Rh(D) chain C03
US-09-240-274-71

Query Match          32.7%; Score 182.2; DB 4; Length 372;
Best Local Similarity 71.3%; Pred. No. 1.6e-49;
Matches 273; Conservative 0; Mismatches 98; Indels 12; Gaps 2;

QY 13 aggtgcagctgcaggagtcagggggagattggtgcaggctggggctctctgagactct 72
      ||||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| |
Db 2 aggtgcagctgctcgagctctgggggagcgctggtccagcctgggaggtccctgagactct 61

QY 73 cctgtgcagcctcgggacgcgccaccagtggtcatggtcactatggtatgggctggttcc 132
      ||||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| |
Db 62 cctgtgcagcctcgtgattctccttcagtag-----ctatggcatgcaactgggtcc 112

QY 133 gccaggttccagggaaggagcgtgagtttctgcagctattaggtggagtggttaaagaga 192
      ||||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| |
Db 113 gccaggtccaggcaagggtggtgagtggtgagttatcatatcatatgatggacatcata 172

QY 193 catggtataaagactccgtgaaggccgattcaccatctccagagataacgccaaagacta 252
      | ||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| |
Db 173 aaaactatgcagactccgtgaaggccgattcaccatctccagagataatccaagaaa 232

QY 253 cggtttatctgcaaatgaacagccctgaaacctgaagatacggccgtttattattgtgccg 312
      || | || ||||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| |
Db 233 cgctgtacctgcaaatgaacagccctgagacctgaggaacacggctgtattactgtgcga 292

QY 313 ctcgaccggtccgcgtggtgatgatatttccctgcgggttgggttgactactg99ggccag 372
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 293 acctaaagggggaagtaactcgtcgtgc---gtctgttcccttgatatcttggggccag 349

QY 373 ggaccaggtcacctgtctctcca 395
      |||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| |
Db 350 ggacaatggtcacgtctcttcca 372

RESULT 14
US-09-240-274-76
; Sequence 76, Application US/09240274
; Patent No. 6255455
; GENERAL INFORMATION:
; APPLICANT: Siegel, Donald L.
; TITLE OF INVENTION: Rh(D)-BINDING PROTEINS AND MAGNETICALLY ACTIVATED CELL
; FILE REFERENCE: 09596-4202
; CURRENT APPLICATION NUMBER: US/09/240,274
; CURRENT FILING DATE: 1999-01-29
; EARLIER FILING DATE: 1998-04-10
; EARLIER APPLICATION NUMBER: 60/081,380
; EARLIER FILING DATE: 1996-10-11
; NUMBER OF SEQ ID NOS: 224
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 76
; LENGTH: 372

Query Match          32.7%; Score 182.2; DB 4; Length 372;
Best Local Similarity 71.3%; Pred. No. 1.6e-49;
Matches 273; Conservative 0; Mismatches 98; Indels 12; Gaps 2;

QY 13 aggtgcagctgcaggagtcagggggagattggtgcaggctggggctctctgagactct 72
      ||||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| |
Db 2 aggtgcagctgctcgagctctgggggagcgctggtccagcctgggaggtccctgagactct 61
```

QY 73 cctgtgcagcctcggaagcgcaccagtggtcactatggtatggctggttcc 132  
|||||  
Db 62 cctgtgcagcctctggtattctcttcagtag-----ctatggcatgcaactgggtcc 112  
QY 133 gccaggttccagggaagcgcgtgagtttgcgcagctattaggtggagtggttaaagaga 192  
|||||  
Db 113 gccaggtccaggcaaggcctggagtggtgtcagttatatcatatgatggacatcata 172  
QY 193 catggtataaagactccgtgaaggccgattccacatctccagagataaacgccaagacta 252  
| |||  
Db 173 aaaactatgcagactccgtgaaggccgattccacatctccagagacaattccaagaaaa 232  
QY 253 cggtttatctgcaaaatgaacagcctgaaacctgaagatacggcgtttattattgtgccg 312  
|| |  
Db 233 cgctgtacctgcaaaatgaacagcctgagacctgaggacacggctgtatatattactgtgcga 292  
QY 313 ctcgaccggtccgcgtgatgatatttccctgcgggttggtttgactactggggccagg 372  
| |  
Db 293 acctaaagggggaagtaactcgtcgtgc---gtctgtccctttgatatactctggtggggccag 349  
QY 373 ggacccaggtcacccgtctctctca 395  
|||||  
Db 350 ggacattggtcacccgtctcttca 372

RESULT 15  
US-09-240-274-79  
; Sequence 79, Application US/09240274  
; Patent No. 6255455  
; GENERAL INFORMATION:  
; APPLICANT: Siegel, Donald L.  
; TITLE OF INVENTION: Rh(D)-BINDING PROTEINS AND MAGNETICALLY ACTIVATED CELL  
; TITLE OF INVENTION: SORTING METHOD FOR PRODUCTION THEREOF  
; FILE REFERENCE: 09596-42U2  
; CURRENT APPLICATION NUMBER: US/09/240,274  
; CURRENT FILING DATE: 1999-01-29  
; EARLIER APPLICATION NUMBER: 60/081,380  
; EARLIER FILING DATE: 1998-04-10  
; EARLIER APPLICATION NUMBER: 60/028,550  
; EARLIER FILING DATE: 1996-10-11  
; NUMBER OF SEQ ID NOS: 224  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 79  
; LENGTH: 378  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: anti-Rh(D) chain D04  
US-09-240-274-79

Query Match 32.6%; Score 182; DB 4; Length 378;  
Best Local Similarity 71.0%; Pred. No. 1.8e-49;  
Matches 274; Conservative 0; Mismatches 100; Indels 12; Gaps 2;  
QY 13 aggtgcagctgcaggagtcagggggaggttggtgcaggctgggggctctctgagactct 72  
|||||  
Db 2 aggtgcagctgctcagtcggtcgggggaggtggtgcccagcctgggaggtccctgagactct 61  
QY 73 cctgtgcagcctcgggacgcgcaccagtggtcatggtcactatggtatggctgggtcc 132  
|||||  
Db 62 cctgtgtagcgtctggattcagctcaggag-----ctatggcatgcactgggtcc 112  
QY 133 gccaggtccagggaaggagcgtgagtttgcgcagctattaggtggagtggttaaagaga 192  
|||||  
Db 113 gccaggtcctggcaaggcctggagtggtggcagatatatggttggaaagtaata 172  
QY 193 catggtataaagactccgtgaaggccgattccacatctccagagataaacgccaagacta 252  
| |||  
Db 173 aagattatgcagactccgtgaaggccgattccacatctccagagacaattccaagaaaca 232  
QY 253 cggtttatctgcaaaatgaacagcctgaaacctgaagatacggccgtttattattgtgc-- 310

Db 233 cgttgtatcttcaaatgaacagcctgagagccgaggatacggcctgtgtattattgtgcga 292  
QY 311 -cgctcgaccggtccgcgtggtgatgatatttccctgcgggttgggtttgactactggggcc 369  
| |  
Db 293 gagattggaggtgcgggaccttagtagtggtggttaagtgccttttgatatctctggggcc 352  
QY 370 aggggacccaggtcacccgtctctctca 395  
| |||||  
Db 353 aagggaacaatggtcacccgtctctctca 378

Search completed: August 13, 2002, 12:31:44  
Job time: 19902 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 13, 2002, 09:24:07 ; Search time 273.29 Seconds  
(without alignments)  
3505.574 Million cell updates/sec

Title: US-09-742-690-1  
Perfect score: 558  
Sequence: 1 tcgagaaaagacaggtgcag.....ctcccagtgctgtgtaataag 558

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_032802: \*  
1: /SIDS5/gcgdata/geneseq/geneseq-emb1/NA1980.DAT: \*  
2: /SIDS5/gcgdata/geneseq/geneseq-emb1/NA1981.DAT: \*  
3: /SIDS5/gcgdata/geneseq/geneseq-emb1/NA1982.DAT: \*  
4: /SIDS5/gcgdata/geneseq/geneseq-emb1/NA1983.DAT: \*  
5: /SIDS5/gcgdata/geneseq/geneseq-emb1/NA1984.DAT: \*  
6: /SIDS5/gcgdata/geneseq/geneseq-emb1/NA1985.DAT: \*  
7: /SIDS5/gcgdata/geneseq/geneseq-emb1/NA1986.DAT: \*  
8: /SIDS5/gcgdata/geneseq/geneseq-emb1/NA1987.DAT: \*  
9: /SIDS5/gcgdata/geneseq/geneseq-emb1/NA1988.DAT: \*  
10: /SIDS5/gcgdata/geneseq/geneseq-emb1/NA1989.DAT: \*  
11: /SIDS5/gcgdata/geneseq/geneseq-emb1/NA1990.DAT: \*  
12: /SIDS5/gcgdata/geneseq/geneseq-emb1/NA1991.DAT: \*  
13: /SIDS5/gcgdata/geneseq/geneseq-emb1/NA1992.DAT: \*  
14: /SIDS5/gcgdata/geneseq/geneseq-emb1/NA1993.DAT: \*  
15: /SIDS5/gcgdata/geneseq/geneseq-emb1/NA1994.DAT: \*  
16: /SIDS5/gcgdata/geneseq/geneseq-emb1/NA1995.DAT: \*  
17: /SIDS5/gcgdata/geneseq/geneseq-emb1/NA1996.DAT: \*  
18: /SIDS5/gcgdata/geneseq/geneseq-emb1/NA1997.DAT: \*  
19: /SIDS5/gcgdata/geneseq/geneseq-emb1/NA1998.DAT: \*  
20: /SIDS5/gcgdata/geneseq/geneseq-emb1/NA1999.DAT: \*  
21: /SIDS5/gcgdata/geneseq/geneseq-emb1/NA2000.DAT: \*  
22: /SIDS5/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT: \*  
23: /SIDS5/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT: \*  
24: /SIDS5/gcgdata/geneseq/geneseq-emb1/NA2002.DAT: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	385.4	69.1	440	AA10052	Llama HCV33 antibo
2	384.8	69.0	1154	AA10057	Llama HCV33 VH reg
3	384.8	69.0	1172	AA10058	Llama HCV33 VH reg
4	299.2	53.6	384	AA59602	PstI/BstEII fragme
5	291.2	52.2	714	AA59612	Plasmid PUR4618 fr
6	291.2	52.2	717	AA59614	Plasmid PUR4620 fr
7	291.2	52.2	750	AA59616	Plasmid PUR4622 fr
8	290.2	52.0	708	AA59617	Plasmid PUR4622 fr
9	232.8	41.7	375	AA78160	Llama anti-carbazo

10	219.8	39.4	378	20	AA78159	Llama anti-carbazo
11	209.6	37.6	354	20	AA59605	PstI/BstEII fragme
12	209.2	37.5	369	20	AA59607	PstI/BstEII fragme
13	208.4	37.3	351	20	AA59604	PstI/BstEII fragme
14	208.4	37.3	351	21	AA00365	Plasmid PUR4601 en
15	208	37.3	462	22	AA10055	Anti-GUS (Clonel8)
16	208	37.3	471	22	AA10056	Anti-GUS VH region
17	201.4	36.1	456	11	AA03610	Sequence encoding
18	199.8	35.8	375	11	AA03608	Sequence encoding
19	198	35.5	867	19	AAV10119	Human H11-scfv con
20	198	35.5	867	22	AA04539	Human monoclonal a
21	197.2	35.3	521	19	AAV17546	Clone CAO5 gene.
22	196.6	35.2	375	22	AAH68638	Human anti-Rh(D) c
23	196.6	35.2	524	19	AAV17545	Clone CAO4 gene.
24	196.4	35.2	867	22	AA04540	Human H11-single c
25	196	35.1	918	19	AAV10118	Human H11-scfv con
26	196	35.1	918	22	AA04537	Human monoclonal a
27	196	35.1	918	22	AA04538	Human H11-single c
28	195.4	35.0	1458	22	AAS22625	Human cDNA encodin
29	195	34.9	375	22	AAH68623	Human anti-Rh(D) c
30	195	34.9	1413	21	AAA46898	DNA encoding the h
31	193.4	34.7	672	20	AA59613	Plasmid PUR4619 fr
32	193.4	34.7	672	21	AA00371	Saccharomyces fusi
33	191.4	34.3	675	20	AA59615	Plasmid PUR4621 fr
34	191	34.2	378	22	AAH68694	Human anti-Rh(D) a
35	190.4	34.1	738	21	AAZ55614	Internalising anti
36	188.6	33.8	372	22	AAF75585	Human anti-HER2/ne
37	188.4	33.8	1741	22	AAS22531	Human cDNA encodin
38	187.6	33.6	681	22	AAH30055	TRO005 heavy chain
39	187.2	33.5	366	22	AAF75589	Human anti-HER2/ne
40	187.2	33.5	1644	22	AAS22593	Human cDNA encodin
41	187	33.5	750	22	AAS00212	DNA encoding monoc
42	187	33.5	753	22	AA00211	DNA encoding monoc
43	187	33.5	909	14	AAQ43772	Sequence encoding
44	185.6	33.3	603	22	AAF85221	Nucleotide sequenc
45	185.4	33.2	381	22	AAH68633	Human anti-Rh(D) c

ALIGNMENTS

RESULT	1
AA10052	
ID	AA10052 standard; DNA; 440 BP.
XX	
AC	AA10052;
XX	
DT	18-SEP-2001 (first entry)
XX	
DE	Llama HCV33 antibody heavy chain variable domain with myc tag DNA.
XX	
KW	Camelid; llama; heavy chain immunoglobulin; pathogen resistance;
KW	metabolism modulator; passive immunisation; heavy chain variable domain;
KW	VH; anti-RR6 antibody; HCV33; azo-dye; ds.
XX	
OS	Chimeric - Lama guanicoe glama.
OS	Chimeric - Unidentified.
XX	
FH	Key Location/Qualifiers
FT	CDS 1..420
FT	/*tag= a
FT	/product= "Llama HCV33 antibody heavy chain variable
FT	domain with myc peptide"
FT	/note= "Does not include start codon"
FT	/partial
FT	
XX	
PN	EP1118669-A2.
XX	
PD	25-JUL-2001.
XX	
PF	08-DEC-2000; 2000EP-0310997.
XX	
PR	17-DEC-1999; 99EP-0310188.





```
CC azo-dye) antibody (designated as HCV33) heavy chain variable domain (VH)
CC linked to the non-classical heavy chain constant regions denoted as
CC hinge-CH2-CH3.
XX
SQ Sequence 1154 BP; 293 A; 341 C; 321 G; 199 T; 0 other;

Query Match          69.0%; Score 384.8; DB 22; Length 1154;
Best Local Similarity 99.5%; Pred. No. 1.2e-95;
Matches 386; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 13 aggtgcagctgcaggagtcaggggaggattggtgcaggctgggggctctctgagactct 72
   |||||||
Db 7 aggtgcagctgcaggagtcaggggaggattggtgcaggctgggggctctctgagactct 66

QY 73 cctgtgcagcctcgggacgcgccaccagtggtcatggtcactatggtatggctggttcc 132
   |||||||
Db 67 cctgtgcagcctcgggacgcgccaccagtggtcatggtcactatggtatggctggttcc 126

QY 133 gccaggttccagggaaggagcgtgagtttgcgcagctattagtggtgagtggttaaagaga 192
   |||||||
Db 127 gccaggttccagggaaggagcgtgagtttgcgcagctattagtggtgagtggttaaagaga 186

QY 193 catggtataaagactccgtgaaggccgattccaccatctccagagataacgcgaagacta 252
   |||||||
Db 187 catggtataaagactccgtgaaggccgattccaccatctccagagataacgcgaagacta 246

QY 253 cggtttatctgcaaatgaacagcgcctgaacctgaagatacaggccggtttatttgtgccg 312
   |||||||
Db 247 cggtttatctgcaaatgaacagcgcctgaacctgaagatacaggccggtttatttgtgccg 306

QY 313 ctgcaccggtccgcgtggatgatatatttccctgcgcggttgggttgactactggtggccagg 372
   |||||||
Db 307 ctgcaccggtccgcgtggatgatatatttccctgcgcggttgggttgactactggtggccagg 366

QY 373 ggaccaggtcacccgtctctcaggatc 400
   |||||||
Db 367 ggaccaggtcacccgtctctcagaacc 394

RESULT 3
AAD10058
ID AAD10058 standard; DNA; 1172 BP.
XX
AC AAD10058;
XX
DT 18-SEP-2001 (first entry)
XX
DE Llama HCV33 VH region with ER retention signal and hinge-CH2-CH3 DNA.
XX
KW Camelid; llama; heavy chain immunoglobulin; pathogen resistance;
KW metabolism modulator; passive immunisation; heavy chain variable domain;
KW VH; anti-RR6 antibody; HCV33; azo-dye; ds.
XX
OS Chimeric - Lama guanicoe glama.
OS Chimeric - Unidentified.
XX
FH Key Location/Qualifiers
FT CDS 3..1154
FT /*tag= a
FT /product= "Llama HCV33 heavy chain variable domain with
FT heavy chain constant regions and ER retention signal"
FT 3..389
FT /*tag= b
FT /note= "Corresponds to Llama HCV33 heavy chain
FT variable domain"
FT 390..476
FT /*tag= c
FT /note= "Corresponds to hinge region DNA"
FT 477..806
FT /*tag= d
FT /note= "Corresponds to heavy chain constant region
FT CH2 DNA"
FT
```

```
FT misc_feature 807..1133
FT /*tag= e
FT /note= "Corresponds to heavy chain constant region
FT CH3 DNA"
FT 1134..1151
FT /*tag= f
FT /note= "Corresponds to ER retention signal DNA"
XX
PN EP1118669-A2.
XX 25-JUL-2001.
XX 08-DEC-2000; 2000EP-0310997.
XX 17-DEC-1999; 99EP-0310188.
XX (UNIL ) UNILEVER PLC.
XX (UNIL ) UNILEVER NV.
XX Frenken LGJ, Van Der Logt CPE, Jobling SA, Teh Y;
PI WPI; 2001-427157/46.
XX P-PSDB; AAE05287.
XX Modifying a plant to produce an antibody useful for increasing pathogen
PT resistance or to modulate metabolism comprises introducing a DNA
PT sequence encoding a heavy chain immunoglobulin linked to a peptide that
PT targets a cellular compartment -
XX Example 12; Fig 19; 81pp; English.
XX The present invention relates to a method for modifying a plant to
CC produce an antibody or an active fragment or derivative, or a protein
CC functional equivalent, in a cellular compartment comprises introducing
CC a DNA sequence encoding a heavy chain immunoglobulin, where the DNA is
CC linked to promoters and provided with an additional sequence encoding a
CC peptide capable of targeting heavy chain immunoglobulin to a cellular
CC compartment. The method is used for producing a heavy chain
CC immunoglobulin or an active fragment or derivative, or a protein that is
CC functionally equivalent for increasing the pathogen resistance in a plant
CC or to modulate metabolism in a plant. Under some circumstances it may be
CC desirable to retain the antibody product with the plant rather than
CC extracting and isolating the product. In particular, edible selected
CC antigens may be used in a method of passively immunising an animal,
CC preferably human, against the antigen, e.g., pathogenic organisms. The
CC present DNA sequence encodes a llama (camelid) anti-RR6 (RR6 is an
CC azo-dye) antibody (designated as HCV33) heavy chain variable domain (VH)
CC linked to the non-classical heavy chain constant regions denoted as
CC hinge-CH2-CH3 and endoplasmic reticulum (ER) retention signal.
XX
SQ Sequence 1172 BP; 300 A; 343 C; 326 G; 203 T; 0 other;

Query Match          69.0%; Score 384.8; DB 22; Length 1172;
Best Local Similarity 99.5%; Pred. No. 1.2e-95;
Matches 386; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 13 aggtgcagctgcaggagtcaggggaggattggtgcaggctgggggctctctgagactct 72
   |||||||
Db 7 aggtgcagctgcaggagtcaggggaggattggtgcaggctgggggctctctgagactct 66

QY 73 cctgtgcagcctcgggacgcgccaccagtggtcatggtcactatggtatggctggttcc 132
   |||||||
Db 67 cctgtgcagcctcgggacgcgccaccagtggtcatggtcactatggtatggctggttcc 126

QY 133 gccaggttccagggaaggagcgtgagtttgcgcagctattagtggtgagtggttaaagaga 192
   |||||||
Db 127 gccaggttccagggaaggagcgtgagtttgcgcagctattagtggtgagtggttaaagaga 186

QY 193 catggtataaagactccgtgaaggccgattccaccatctccagagataacgcgaagacta 252
   |||||||
Db 187 catggtataaagactccgtgaaggccgattccaccatctccagagataacgcgaagacta 246
```

QY 253 cgggttatctgcaaatgaacagccctgaaacccctgaagatacggccggtttattattgtgccc 312  
Db 247 cgggttatctgcaaatgaacagccctgaaacccctgaagatacggccggtttattattgtgccc 306  
QY 313 ctgcgaccggtccgcgtgatgatatttccctgcgcggttggttgactactcggggccagg 372  
Db 307 ctgcgaccggtccgcgtgatgatatttccctgcgcggttggttgactactcggggccagg 366  
QY 373 ggaccaggtcacccgtctctcaggatc 400  
Db 367 ggaccaggtcacccgtctctcagaacc 394  
RESULT 4  
AAX59602  
ID AAX59602 standard; DNA: 384 BP.  
XX  
AC AAX59602;  
XX  
DT 21-JUL-1999 (first entry)  
XX  
DE PstI/BstEII fragment of plasmid pUR4538.  
XX  
KW Multivalent antigen binding protein; single polypeptide chain;  
KW single domain binding site; diagnosis; therapy; targeting;  
KW immunoassay; cross-linking; agglutination; purification;  
KW phage inactivation; detergent; ss.  
XX  
OS Synthetic.  
XX  
PN WO9923221-A2.  
XX  
PD 14-MAY-1999.  
XX  
PF 27-OCT-1998; 98WO-EP06991.  
XX  
PR 27-OCT-1997; 97EP-0308538.  
XX  
PA (UNIL ) UNILEVER NV.  
PA (UNIL ) UNILEVER PLC.  
XX  
PI Frenken LGJ, Howell S, Ledebroer AM, Van Der Logt CPE;  
XX WPI; 1999-313342/26.  
DR P-PSDB; AAY15398.  
XX  
PT Multivalent antigen binding protein useful for inactivation of  
PT (bacterio)phages or viruses  
XX  
PS Disclosure; Fig 2; 89pp; English.  
XX  
CC The specification describes multivalent antigen binding proteins  
CC comprising a single polypeptide chain comprising, in series, two or  
CC more single domain binding sites. The multivalent antigen proteins  
CC are useful in applications where antibodies have been used in the  
CC prior art, including diagnosis, therapy, targeting, immunoassays,  
CC cross-linking methods including agglutination, or for purification  
CC processes. It is also useful for inactivation of (bacterio)phages  
CC or viruses, and in detergents. The present sequence encodes the  
CC heavy chain variable domain of an anti-RR6 antibody from a llama.  
XX  
SQ Sequence 384 BP; 78 A; 96 C; 124 G; 86 T; 0 other;  
Query Match 53.6%; Score 299.2; DB 20; Length 384;  
Best Local Similarity 86.2%; Pred. No. 2.5e-72;  
Matches 331; Conservative 0; Mismatches 53; Indels 0; Gaps 0;  
QY 12 cagggtgcagctgcaggagtcaggggaggaggttggtgcaggctgggggtctctcaggactc 71  
Db 1 cagggtgcagctgcaggagtcaggggaggaggttggtgcaggctgggggtctctcaggactc 60  
QY 72 tctgtgcagcctcgggacgcgcaccagtggtcattggtcactatggtatgggctggttc 131

Db 61 tctgtcgcgcctcgggacgcacttctcatggttatggttgctatggcatgggctggttc 120  
QY 132 cgccaggttccagggaaggagcgtgagttgttcgacgtattaggtggagtggtaaagag 191  
Db 121 cgccaaattccagggaaggagcgtgagcttgcgagcaattaggtggagcgtcgtaaat 180  
QY 192 acatggtataaagactccgtgaaggccgattccaccatctccagagataacgcacaagact 251  
Db 181 acatactatgcagactccgtgaaggccgattccaccatctccagagacaacgtcaaggac 240  
QY 252 acggtttatctgcaaatgaacagccctgaaacctgaagatacggccgtttattattgtgcc 311  
Db 241 atgctgtatctgcaaatgaacagtttgaaacctgaggacacggccgtttacacttgca 300  
QY 312 gctcgaccggtccgcgtgatgatatttccctgcgcggttggttgactactcggggccag 371  
Db 301 gttcggacggtccgcgtggttgacatttccagtcgcggttggttgctactcggggccag 360  
QY 372 gggaccaggtcacccgtctctcctca 395  
Db 361 gggaccaggtcacccgtctctcctca 384  
RESULT 5  
AAX59612  
ID AAX59612 standard; DNA: 714 BP.  
XX  
AC AAX59612;  
XX  
DT 21-JUL-1999 (first entry)  
XX  
DE Plasmid pUR4618 fragment encoding an anti-hCG anti-RR6 ABP.  
XX  
KW Multivalent antigen binding protein; single polypeptide chain;  
KW single domain binding site; diagnosis; therapy; targeting;  
KW immunoassay; cross-linking; agglutination; purification;  
KW phage inactivation; detergent; ss.  
XX  
OS Synthetic.  
XX  
PN WO9923221-A2.  
XX  
PD 14-MAY-1999.  
XX  
PF 27-OCT-1998; 98WO-EP06991.  
XX  
PR 27-OCT-1997; 97EP-0308538.  
XX  
PA (UNIL ) UNILEVER NV.  
PA (UNIL ) UNILEVER PLC.  
XX  
PI Frenken LGJ, Howell S, Ledebroer AM, Van Der Logt CPE;  
XX WPI; 1999-313342/26.  
PT Multivalent antigen binding protein useful for inactivation of  
PT (bacterio)phages or viruses  
XX  
PS Disclosure; Fig 27; 89pp; English.  
XX  
CC The specification describes multivalent antigen binding proteins  
CC comprising a single polypeptide chain comprising, in series, two or  
CC more single domain binding sites. The multivalent antigen proteins  
CC are useful in applications where antibodies have been used in the  
CC prior art, including diagnosis, therapy, targeting, immunoassays,  
CC cross-linking methods including agglutination, or for purification  
CC processes. It is also useful for inactivation of (bacterio)phages  
CC or viruses, and in detergents. The present sequence encodes an  
CC anti-hCG anti-RR6 bispecific biheaded antigen binding protein (ABP).  
XX  
SQ Sequence 714 BP; 149 A; 181 C; 237 G; 147 T; 0 other;

Query Match 52.2%; Score 291.2; DB 20; Length 714;  
Best Local Similarity 85.9%; Pred. No. 4.5e-70;  
Matches 323; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

QY 11 acaggtgcagctgcaggagtcaggggagagctggtgagctggtgagctggtgagact 70  
Db 339 acaggtgcagctgcaggagtcaggggagagctggtgagctggtgagact 398

QY 71 ctctgtgcagctgcgggacgcgcacacagtggtcactggtcactggtggtggtggt 130  
Db 399 ctctgcgcgctgcgggacgcgcacttctcactggtggtggtggtggtggtggt 458

QY 131 ccgccaggttcagggaaggagcgtgagctggtggtggtggtggtggtggtggtggt 190  
Db 459 ccgccaattccagggaaggagcgtgagctggtggtggtggtggtggtggtggtggt 518

QY 191 gacatggtataaagactccgtgaaggccgacttccaccatctccagagataacgcgaagac 250  
Db 519 tacatactatgcagactccgtgaaggccgacttccaccatctccagagataacgcgaagga 578

QY 251 tacggtttatctgcaaatgaacagcctgaaacctgaagatacagcgcgtttattattgtgc 310  
Db 579 catgctgtatctgcaaatgaacagcttgaacctgaggacacgcgcgtttacacttgc 638

QY 311 cgctcgaccggtccgctggtgatgatatttccctgcgcggttgggtttgactactggtggcca 370  
Db 639 agttcggacggtccgctggttgacatttccagttccagtcgcggttgggtttgctactggtggcca 698

QY 371 ggggaccaggtcaacc 386  
Db 699 ggggaccaggtcaacc 714

RESULT 6  
AA59614  
ID AAX59614 standard; DNA; 717 BP.  
XX  
AC AAX59614;  
XX  
DT 21-JUL-1999 (first entry)  
XX  
DE Plasmid PUR4620 fragment encoding an anti-hCG anti-RR6 ABP.

XX Multivalent antigen binding protein; single polypeptide chain;  
KW single domain binding site; diagnosis; therapy; targeting;  
KW immunoassay; cross-linking; agglutination; purification;  
KW phage inactivation; detergent; ss.  
XX  
OS Synthetic.  
XX  
PN WO9923221-A2.  
XX  
PD 14-MAY-1999.  
XX  
PF 27-OCT-1998; 98WO-EP06991.  
XX  
PR 27-OCT-1997; 97EP-0308538.  
XX  
PA (UNIL ) UNILEVER NV.  
PA (UNIL ) UNILEVER PLC.  
XX  
PI Frenken LGJ, Howell S, Ledebøer AM, Van Der Logt CPE;  
XX  
DR WPI; 1999-313342/26.  
DR P-PSDB; AAY15408.  
XX  
PT Multivalent antigen binding protein useful for inactivation of  
PT (bacterio)phages or viruses  
XX  
PS Disclosure; Fig 10; 89pp; English.  
XX  
CC The specification describes multivalent antigen binding proteins

CC comprising a single polypeptide chain comprising, in series, two or  
CC more single domain binding sites. The multivalent antigen proteins  
CC are useful in applications where antibodies have been used in the  
CC prior art, including diagnosis, therapy, targeting, immunoassays,  
CC cross-linking methods including agglutination, or for purification  
CC processes. It is also useful for inactivation of (bacterio)phages  
CC or viruses, and in detergents. The present sequence encodes an  
CC anti-hCG anti-RR6 bispecific biheaded antigen binding protein (ABP).  
XX

SQ Sequence 717 BP; 158 A; 183 C; 220 G; 156 T; 0 other;

Query Match 52.2%; Score 291.2; DB 20; Length 717;  
Best Local Similarity 85.9%; Pred. No. 4.5e-70;  
Matches 323; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

QY 11 acaggtgcagctgcaggagtcaggggagagctggtgagctggtggtggtggtggtggt 70  
Db 342 acaggtgcagctgcaggagtcaggggagagctggtgagctggtggtggtggtggtggt 401

QY 71 ctctgtgcagctgcgggacgcgcacacagtggtcactggtcactggtggtggtggtt 130  
Db 402 ctctgcgcgctgcgggacgcgcacttctcactggtggtggtggtggtggtggtggtt 461

QY 131 ccgccaggttcagggaaggagcgtgagctggtggtggtggtggtggtggtggtggt 190  
Db 462 ccgccaattccagggaaggagcgtgagctggtggtggtggtggtggtggtggtggt 521

QY 191 gacatggtataaagactccgtgaaggccgacttccaccatctccagagataacgcgaagac 250  
Db 522 tacatactatgcagactccgtgaaggccgacttccaccatctccagagataacgcgaagga 581

QY 251 tacggtttatctgcaaatgaacagcctgaaacctgaagatacagcgcgtttattattgtgc 310  
Db 582 catgctgtatctgcaaatgaacagcttgaacctgaggacacgcgcgtttacacttgc 641

QY 311 cgctcgaccggtccgctggtgatgatatttccctgcgcggttgggtttgactactggtggcca 370  
Db 642 agttcggacggtccgctggttgacatttccagttccagtcgcggttgggtttgctactggtggcca 701

QY 371 ggggaccaggtcaacc 386  
Db 702 ggggaccaggtcaacc 717

RESULT 7  
AA59616  
ID AAX59616 standard; DNA; 750 BP.  
XX  
AC AAX59616;  
XX  
DT 21-JUL-1999 (first entry)  
XX  
DE Plasmid PUR4622 fragment encoding an anti-RR6 antigen binding protein.  
XX  
KW Multivalent antigen binding protein; single polypeptide chain;  
KW single domain binding site; diagnosis; therapy; targeting;  
KW immunoassay; cross-linking; agglutination; purification;  
KW phage inactivation; detergent; ss.  
XX  
OS Synthetic.  
XX  
PN WO9923221-A2.  
XX  
PD 14-MAY-1999.  
XX  
PF 27-OCT-1998; 98WO-EP06991.  
XX  
PR 27-OCT-1997; 97EP-0308538.  
XX  
PA (UNIL ) UNILEVER NV.  
PA (UNIL ) UNILEVER PLC.  
XX



PI Frenken LGJ, Howell S, Ledebroer AM, Van Der Logt CPE;  
XX  
DR WPI; 1999-313342/26.  
DR P-PSDB; AAY15413.  
XX  
PT Multivalent antigen binding protein useful for inactivation of  
PT (bacterio)phages or viruses  
XX  
PS Disclosure; Fig 12; 89pp; English.  
XX  
XX The specification describes multivalent antigen binding proteins  
CC comprising a single polypeptide chain comprising, in series, two or  
CC more single domain binding sites. The multivalent antigen proteins  
CC are useful in applications where antibodies have been used in the  
CC prior art, including diagnosis, therapy, targeting, immunoassays,  
CC cross-linking methods including agglutination, or for purification  
CC processes. It is also useful for inactivation of (bacterio)phages  
CC or viruses, and in detergents. The present sequence encodes an  
CC homodimeric bivalent anti-RR6 antigen binding protein.  
XX  
SQ Sequence 750 BP; 153 A; 186 C; 243 G; 168 T; 0 other;

Query Match 52.2%; Score 291.2; DB 20; Length 750;  
Best Local Similarity 85.9%; Pred. No. 4.5e-70;  
Matches 323; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

QY 11 acaggtgcagctgcaggagtcaggggagggattggtgcaggctggggctctctgagact 70  
Db 375 acaggtgcagctgcaggagtcaggggagggattggtgcaggctggggactctctgagact 434  
QY 71 ctctgtgcagcctcgggacgcgcaccagtggtcactatggtcattggtggctggtt 130  
Db 435 ctctgcgcggcctcgggacgcacttctcactatggtggtatggtgcattggtggtt 494  
QY 131 ccgccaggttccaggaaggagcgtgagtttgcagctattagtggtgagtggtaaaga 190  
Db 495 ccgccaaattccaggaaggagcgtgagcttgcagcaattagtggtgagcgtcgtaa 554  
QY 191 gacatggtataaaagactccgtgaaggccgattaccatctccagagataacgccaaagac 250  
Db 555 tacatactatgcagactccgtgaaggccgattaccatctccagagataacgcagga 614  
QY 251 tacggttatctgcaaatgaacagccctgaaacctgaagatacggcgtttattattgtgc 310  
Db 615 catgctgtatctgcaaatgaacagtttgaaacctgaggacacggcgtttacactgtgc 674  
QY 311 cgctcgaccggtccgctggtgatgatatttccctgcgggttgggttgactactcaggccca 370  
Db 675 agttcggacggtccgctggttgacatttccagtcgggttgggttgactcaggccca 734  
QY 371 ggggacccaggtcacc 386  
Db 735 ggggacccaggtcacc 750

RESULT 8  
AAX59617  
ID AAX59617 standard; DNA; 708 BP.  
XX  
AC AAX59617;  
XX  
DT 21-JUL-1999 (first entry)  
XX  
DE Plasmid PUR4622 fragment encoding an anti-RR6 antigen binding protein.  
XX  
KW Multivalent antigen binding protein; single polypeptide chain;  
KW single domain binding site; diagnosis; therapy; targeting;  
KW immunoassay; cross-linking; agglutination; purification;  
KW phage inactivation; detergent; ss.  
XX  
OS Synthetic.  
XX

PN WO9923221-A2.  
XX  
PD 14-MAY-1999.  
XX  
PF 27-OCT-1998; 98WO-EP06991.  
XX  
PR 27-OCT-1997; 97EP-0308538.  
XX  
PA (UNIL ) UNILEVER NV.  
PA (UNIL ) UNILEVER PLC.  
XX  
PI Frenken LGJ, Howell S, Ledebroer AM, Van Der Logt CPE;  
XX  
DR WPI; 1999-313342/26.  
DR P-PSDB; AAY15414.  
XX  
PT Multivalent antigen binding protein useful for inactivation of  
PT (bacterio)phages or viruses  
XX  
PS Disclosure; Fig 13; 89pp; English.  
XX  
CC The specification describes multivalent antigen binding proteins  
CC comprising a single polypeptide chain comprising, in series, two or  
CC more single domain binding sites. The multivalent antigen proteins  
CC are useful in applications where antibodies have been used in the  
CC prior art, including diagnosis, therapy, targeting, immunoassays,  
CC cross-linking methods including agglutination, or for purification  
CC processes. It is also useful for inactivation of (bacterio)phages  
CC or viruses, and in detergents. The present sequence encodes an  
CC heterodimeric bivalent anti-RR6 antigen binding protein.  
XX  
SQ Sequence 708 BP; 159 A; 178 C; 217 G; 154 T; 0 other;

Query Match 52.0%; Score 290.2; DB 20; Length 708;  
Best Local Similarity 85.9%; Pred. No. 8.3e-70;  
Matches 322; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

QY 21 ctgcaggagtcaggggaggttggtgcaggctgggggctctctgagactctcctgtgca 80  
Db 1 ctgcaggagtcaggggaggttggtgcaggctgggggactctctgagactctcctgcgcg 60  
QY 81 gcctcgggacgcgccaccagtggtcatggtcactatggtatggctggttccgccaggtt 140  
Db 61 gcctcgggacgcacttctcatggtggtatggtggtatggctggttccgccaaatt 120  
QY 141 ccagggaaggagcgtgagtttgcgcagctattagtggtggttaaagagacatggtat 200  
Db 121 ccagggaaggagcgtgagcttgcgcagcaattagtggtgagcgtcgtataatactat 180  
QY 201 aaagactccgtgaaggcccgattcaccatctccagagataaacgcaagactcggtttat 260  
Db 181 gcagactccgtgaaggcccgattcaccatctccagagataaacgcaagactcgtgtat 240  
QY 261 ctgcaaatgaacagcctgaaacctgaagatacggcgtttattattgtgccgctcgaccg 320  
Db 241 ctgcaaatgaacagtttgaaacctgaggacacggccgtttacactgtgcagttcggacg 300  
QY 321 gtccgcgtggtatgatatttccctgcgggttgggttgactactcaggccaggggacccag 380  
Db 301 gtccgcgtggttgacatttccagtcgggttgggttgcctactcaggccaggggacccag 360  
QY 381 gtcaccgtctcctca 395  
Db 361 gtcaccgtctcctca 375

RESULT 9  
AAX78160  
ID AAX78160 standard; DNA; 375 BP.  
XX  
AC AAX78160;  
XX

DT 20-AUG-1999 (first entry)  
DE Llama anti-carbazole HC-V fragment 24 DNA.  
XX  
KW Analytical device; electrically conductive solid support; immobilisation;  
KW electroactive; analysis; peptide hormone; steroid; disease marker;  
KW diagnostic indicator; antibody; bacteria; virus; fungi; yeast; llama;  
KW HCV-24; ss.  
XX  
OS Lama guanicoe glama.  
XX  
PN WO9927356-A1.  
XX  
PD 03-JUN-1999.  
XX  
PF 23-NOV-1998; 98WO-GB03495.  
XX  
PR 21-NOV-1997; 97EP-0309425.  
XX  
PA (UNIL ) UNILEVER NV.  
PA (UNIL ) UNILEVER PLC.  
XX  
PI Badley RA, Porter RA;  
XX  
XX WPI; 1999-385228/32.  
DR P-PSDB; AAY08985.  
XX  
XX Analytical device including electrochemically active compound on solid support  
PT  
PT  
XX  
PS Example 6.2; Page 66; 78pp; English.  
XX  
CC This invention describes a novel component for an analytical device comprising an electrically conductive solid support on which a compound (I) is immobilized which has an electroactive part (EA) with an electrochemical property that can be modulated in a detectable manner by binding it to a specific binding partner (II). Compounds used in the method of the invention include those with the formula R = hydrogen, hydroxy, 1-14C alkyl, aryl, alkenyl or alkoxy (all optionally substituted), halo, amido, or amino; optionally one or more positions on the heteroaromatic rings are substituted by alkyl, aryl, alkenyl or alkoxy (all optionally substituted), acid groups (organic or inorganic), halo, amido or amino. Devices that contain the component of the invention are used for qualitative and quantitative analysis of e.g. nucleic acid, hormones (peptide or steroid), disease markers, diagnostic indicators etc., e.g. antibodies, bacteria, viruses, fungi or yeast. These devices can be used with turbid solutions (whole blood, milk, culture liquids etc.), unlike conventional colourimetric methods. This sequence encodes a llama HC-V fragment, HCV24 which is used in the method of the invention.  
XX  
SQ Sequence 375 BP; 83 A; 90 C; 117 G; 85 T; 0 other;

Query Match 41.7%; Score 232.8; DB 20; Length 375;  
Best Local Similarity 77.6%; Pred. No. 3.6e-54;  
Matches 298; Conservative 0; Mismatches 77; Indels 9; Gaps 1;

QY 12 caggtgcagctgcaggagtcaggggaggattggtgcaggctgggggctctctgagactc 71  
|||||  
Db 1 caggtgcagctgcaggagtcaggggaggattggtgcaggctgggggctctctgagactc 60  
QY 72 tcctgtgcagcctgggacgcgccaccagtggtggtcactatggtatgggtggttc 131  
|||||  
Db 61 tcctgtgcagcctgggacgcgccacctcagtt-----tatgccgtgggtggttc 111  
QY 132 cgccaggttcagggaagcgtgagtttgcagctattaggtgagtggttaagag 191  
|||||  
Db 112 cgccaggttcagggaagcgtgagtttgcagctattggtgagcgtggtggaaga 171  
QY 192 acatggtataaagactccgtgaaggccgattaccatctccagagataacccaagact 251  
|||||  
Db 172 acatactatgcagactccgtgaaggccgattaccatctccagagataacccaagac 231

QY 252 acggtttatctgcaaatgaacagcctgaaacccctgaacatacggccgtttattattgtgcc 311  
|||||  
Db 232 acggtgtatctgcaaatgaatagcctgaaactagacgatacggccgtttattactgcgca 291  
QY 312 gctcgaccggtccgctggtgatgatatttccctgcoggttggttgactactggggccag 371  
|||||  
Db 292 gtccgtatgcogtatagtggtgattaccgcatctagtgggacatatgactactggggccag 351  
QY 372 gggacccaggtcacctctcctca 395  
|||||  
Db 352 gggacccaggtcacctctcctca 375  
RESULT 10  
AAX78159  
ID AAX78159 standard; DNA; 378 BP.  
XX  
AC AAX78159;  
XX  
DT 20-AUG-1999 (first entry)  
XX  
DE Llama anti-carbazole HC-V fragment 3 DNA.  
XX  
KW Analytical device; electrically conductive solid support; immobilisation;  
KW electroactive; analysis; peptide hormone; steroid; disease marker;  
KW diagnostic indicator; antibody; bacteria; virus; fungi; yeast; llama;  
KW HCV-3; ss.  
XX  
OS Lama guanicoe glama.  
XX  
PN WO9927356-A1.  
XX  
PD 03-JUN-1999.  
XX  
PF 23-NOV-1998; 98WO-GB03495.  
XX  
PR 21-NOV-1997; 97EP-0309425.  
XX  
PA (UNIL ) UNILEVER NV.  
PA (UNIL ) UNILEVER PLC.  
XX  
PI Badley RA, Porter RA;  
XX  
XX WPI; 1999-385228/32.  
DR P-PSDB; AAY08984.  
XX  
PT Analytical device including electrochemically active compound on solid support  
XX  
PS Example 6.2; Page 65; 78pp; English.  
XX  
CC This invention describes a novel component for an analytical device comprising an electrically conductive solid support on which a compound (I) is immobilized which has an electroactive part (EA) with an electrochemical property that can be modulated in a detectable manner by binding it to a specific binding partner (II). Compounds used in the method of the invention include those with the formula R = hydrogen, hydroxy, 1-14C alkyl, aryl, alkenyl or alkoxy (all optionally substituted), halo, amido, or amino; optionally one or more positions on the heteroaromatic rings are substituted by alkyl, aryl, alkenyl or alkoxy (all optionally substituted), acid groups (organic or inorganic), halo, amido or amino. Devices that contain the component of the invention are used for qualitative and quantitative analysis of e.g. nucleic acid, hormones (peptide or steroid), disease markers, diagnostic indicators etc., e.g. antibodies, bacteria, viruses, fungi or yeast. These devices can be used with turbid solutions (whole blood, milk, culture liquids etc.), unlike conventional colourimetric methods. This sequence encodes a llama HC-V fragment, HCV3 which is used in the method of the invention.  
XX  
SQ Sequence 378 BP; 89 A; 94 C; 112 G; 83 T; 0 other;

Query Match 39.4%; Score 219.8; DB 20; Length 378;

Best Local Similarity 77.0%; Pred. No. 1.3e-50; Mismatches 77; Indels 12; Gaps 2; Matches 298; Conservative 0;

Qy 12 caggtgcagctgcaggagtcaggaggagattggtgcaggtcggggctctctgagactc 71  
|||||  
Db 1 caggtgcagctgcaggagtcaggaggagattggtgcaggtcggggctctctgagactc 60  
|||||

Qy 72 tcctgtgcagctcgggacgcgcaccagtggtcatggtcactatggtatggctggttc 131  
|||||  
Db 61 tcctgtgcagcttctggactcacattga-----ctacctattcaacgggctggttc 111  
|||||

Qy 132 cgccaggttccagggaaggagcgtgagtttgcgcagctattaggtggagtggtaaaag-- 189  
|||||  
Db 112 cgccaggttccagggaaggagcgtgagtttgcgcagctattaggtggagtgggtggc 171  
|||||

Qy 190 -agacatggtataaaagactccgtgaaggccgattcaccatctccagagataacgccaaag 248  
|||||  
Db 172 aacacgtactacgcagactccgtgaaggccgatttaccatctccagagacaaagccaaag 231  
|||||

Qy 249 actacggtttatctgcaaaatgaacagccctgaaacctgaagatacggccgtttattattgt 308  
|||||  
Db 232 aatatggtgttctgcaaaatgagcagcctgaaacctgaggacacggccgtttattactgt 291  
|||||

Qy 309 gccgtgcagcgggtccgcgtggatgatatttccctgcgggttggttgactactggggc 368  
|||||  
Db 292 gcagcagcacaacccctaccgaggtagttacagtgatcccgaaataattatcattactggggc 351  
|||||

Qy 369 caggggacccaggtcacccgtctcctca 395  
|||||  
Db 352 caggggacccaggtcacccgtctcctca 378  
|||||

RESULT 11  
AAx59605  
ID AAX59605 standard; DNA; 354 BP.  
XX  
AC AAX59605;  
XX  
DT 21-JUL-1999 (first entry)  
XX  
DE PstI/BstEII fragment of plasmid PUR4602.  
XX  
KW Multivalent antigen binding protein; single polypeptide chain;  
KW single domain binding site; diagnosis; therapy; targeting;  
KW immunoassay; cross-linking; agglutination; purification;  
KW phage inactivation; detergent; ss.  
XX  
OS Synthetic.  
XX  
PN WO9923221-A2.  
XX  
PD 14-MAY-1999.  
XX  
PF 27-OCT-1998; 98WO-EP06991.  
XX  
PR 27-OCT-1997; 97EP-0308538.  
XX  
PA (UNIL ) UNILEVER NV.  
PA (UNIL ) UNILEVER PLC.  
XX  
PI Frenken LGJ, Howell S, Ledebøer AM, Van Der Logt CPE;  
XX  
DR WPI; 1999-313342/26.  
DR P-PSDB; AAY15401.  
XX  
PT Multivalent antigen binding protein useful for inactivation of  
PT (bacterio)phages or viruses  
XX  
PS Disclosure; Fig 5; 89pp; English.  
XX  
CC The specification describes multivalent antigen binding proteins  
CC comprising a single polypeptide chain comprising, in series, two or  
CC more single domain binding sites. The multivalent antigen proteins

are useful in applications where antibodies have been used in the prior art, including diagnosis, therapy, targeting, immunoassays, cross-linking methods including agglutination, or for purification processes. It is also useful for inactivation of (bacterio)phages or viruses, and in detergents. The present sequence encodes the heavy chain variable domain of an anti-hCG antibody HI-15 from a llama.

XX  
SQ Sequence 354 BP; 84 A; 93 C; 103 G; 74 T; 0 other;

Query Match 37.6%; Score 209.6; DB 20; Length 354;  
Best Local Similarity 76.8%; Pred. No. 8e-48;  
Matches 295; Conservative 0; Mismatches 59; Indels 30; Gaps 2;

Qy 12 caggtgcagctgcaggagtcaggaggagattggtgcaggtcggggctctctgagactc 71  
|||||  
Db 1 caggtgcagctgcaggagtcgtctggggagaattggtgcagcctggggctctctgaaactc 60  
|||||

Qy 72 tcctgtgcagctcgggacgcgcaccagtggtcatggtcactatggtatggctggttc 131  
|||||  
Db 61 tcctgcgcagcctctggacttaccttca-----ctaattatagcatggctggttc 111  
|||||

Qy 132 cgccaggttccagggaaggagcgtgagtttgcgcagctattaggtggagtggtaaaagag 191  
|||||  
Db 112 cgccaggttccaggagtggtgacgtgagccgtagccgtattagctggagtggatgataac 171  
|||||

Qy 192 acatggtataaaagactccgtgaaggccgattcaccatctccagagataacgccaaagact 251  
|||||  
Db 172 acatactatgtaagctccgtgaaggacgattcaccatctccagagacaaagcaaac 231  
|||||

Qy 252 acggtttatctgcaaaatgaacagcctgaaacctgaagatacagccgtttattattgtgcc 311  
|||||  
Db 232 acggtgtatctgcaaaatgaacagcctgaaacctgaagatacagccgtttattattgtgca 291  
|||||

Qy 312 gctcgaccgtccgcgtggatgatatttccctgcgggttggttgactactggggccag 371  
|||||  
Db 292 gtaaaacccgac-----gatggtggtgggactactggggccag 330  
|||||

Qy 372 gggacccaggtcacccgtctcctca 395  
|||||  
Db 331 gggacccaggtcacccgtctcctca 354  
|||||

RESULT 12  
AAx59607  
ID AAX59607 standard; DNA; 369 BP.  
XX  
AC AAX59607;  
XX  
DT 21-JUL-1999 (first entry)  
XX  
DE PstI/BstEII fragment of plasmid PUR4642.  
XX  
KW Multivalent antigen binding protein; single polypeptide chain;  
KW single domain binding site; diagnosis; therapy; targeting;  
KW immunoassay; cross-linking; agglutination; purification;  
KW phage inactivation; detergent; ss.  
XX  
OS Synthetic.  
XX  
PN WO9923221-A2.  
XX  
PD 14-MAY-1999.  
XX  
PF 27-OCT-1998; 98WO-EP06991.  
XX  
PR 27-OCT-1997; 97EP-0308538.  
XX  
PA (UNIL ) UNILEVER NV.  
PA (UNIL ) UNILEVER PLC.  
XX  
PI Frenken LGJ, Howell S, Ledebøer AM, Van Der Logt CPE;  
XX





KW linker; conformational flexibility; multivalent binding protein;  
KW bi-head; immunoassay; agglutination assay; purification; antibody;  
KW human chorionic gonadotropin; hCG; ds.  
XX  
OS Lama guanicoe glama.  
OS Homo sapiens.  
OS Synthetic.  
XX  
FH Key Location/Qualifiers  
FT CDS 1..351  
FT /\*tag= a  
FT /product= "Llama HC-V domain of anti-hCG antibody (H14)"  
FT /note= "Does not include stop codon"  
FT /partial  
XX  
PN WO200024884-A2.  
XX  
XX 04-MAY-2000.  
PD  
XX 22-OCT-1999; 99WO-EP08323.  
PF  
XX 27-OCT-1998; 98WO-EP06991.  
PR 22-APR-1999; 99EP-0303118.  
PR  
XX (UNIL ) UNILEVER PLC.  
PA (UNIL ) UNILEVER NV.  
PA (HIND-) HINDUSTAN LEVER LTD.  
XX  
PI Frenken LGJ, Howell S, Van Der Vaart JM;  
XX  
XX WPI; 2000-350728/30.  
DR P-PSDB; AAY70992.  
DR  
XX Use of a linker whose amino acid sequence confers restricted  
PT conformational flexibility to generate multivalent and multispecific  
PT antigen binding proteins -  
XX  
PS Example 1.1b; Fig 2; 50pp; English.  
XX  
CC The present DNA sequence is the PstI-BstEII insert of the plasmid  
CC PUR4601, encoding llama heavy chain variable (HC-V) domain of an anti-hCG  
CC antibody (denoted as H14). The PUR4601 plasmid comprising this sequence  
CC is transformed into Saccharomyces cerevisiae. They are screened for  
CC antigen specific HC-V fragments that specifically bind to human chorionic  
CC gonadotropin (hCG) antigen. This plasmid is used for the construction of  
CC antigen binding proteins, comprising two single binding units referred to  
CC as Bi-heads with various linkers. The peptide linker confers restricted  
CC conformational flexibility for linking binding units in a multivalent  
CC binding protein. The linker is used to generate multivalent or  
CC multispecific antigen binding proteins for immunoassays, agglutination  
CC assays or for purification.  
XX  
SQ Sequence 351 BP; 75 A; 91 C; 120 G; 65 T; 0 other;

Query Match 37.3%; Score 208.4; DB 21; Length 351;  
Best Local Similarity 83.4%; Pred. No. 1.7e-47;  
Matches 252; Conservative 0; Mismatches 41; Indels 9; Gaps 1;

Qy 12 cagggtcagctgcaggagtcaggaggaggttggtgcaggctgggggctctctgagactc 71  
|||||  
Db 1 cagggtcagctgcaggagtcaggaggaggttggtgcaggcgggggctctctgagactc 60  
|||||  
Qy 72 tcctgtgcagcctcgggacgcgcaccagtggtcatggtcactatggtatggctggttc 131  
|||||  
Db 61 tcctgtgcagcctctggacgcaccggca-----gtacgtatgacatgggctggttc 111  
|||||  
Qy 132 cgccaggttcagggaaggagcgtgagttgtcgcagctattagtggtggttaaagag 191  
|||||  
Db 112 cgccaggttcagggaaggagcgtgagttgtcgcagctattaaactgggtagtgcgcgc 171  
|||||  
Qy 192 acatgggtataaagactccgtgaaggccgattcaccatctccagagataacgccaaagact 251  
|||||

Db 172 acatactatgcaagctccgtgaggggcgcgattcaccatctccagagacaacgccaagaag 231  
Qy 252 acggtttatctgcaaatgaacagcctgaaacctgaagatacggccgtttattattgtgcc 311  
|||||  
Db 232 acggtgtatctgcaaatgaacagcctgaaacctgaggacacggccgtttatacctgtggc 291  
Qy 312 gc 313  
||  
Db 292 gc 293  
||  
RESULT 15  
AAD10055  
ID AAD10055 standard; DNA; 462 BP.  
XX  
AC AAD10055;  
XX  
DT 18-SEP-2001 (first entry)  
XX  
DE Anti-GUS (Clone18) VH region attached with myc and his6 tag DNA.  
XX  
KW Heavy chain immunoglobulin; pathogen resistance; metabolism modulator;  
KW passive immunisation; heavy chain variable domain; VH; anti-GUS antibody;  
KW beta-glucuronidase; ds.  
XX  
OS Unidentified.  
XX  
FH Key Location/Qualifiers  
FT CDS 1..453  
FT /\*tag= a  
FT /product= "Anti-GUS heavy chain variable domain  
FT with myc and his6 peptide"  
FT /note= "Does not include start codon"  
FT /partial  
XX  
PN EP1118669-A2.  
XX  
PD 25-JUL-2001.  
XX  
PF 08-DEC-2000; 2000EP-0310997.  
XX  
PR 17-DEC-1999; 99EP-0310188.  
XX  
PA (UNIL ) UNILEVER PLC.  
PA (UNIL ) UNILEVER NV.  
XX  
PI Frenken LGJ, Van Der Logt CPE, Jobling SA, Teh Y;  
XX  
XX WPI; 2001-427157/46.  
DR P-PSDB; AAE05284.  
DR  
XX  
PT Modifying a plant to produce an antibody useful for increasing pathogen  
PT resistance or to modulate metabolism comprises introducing a DNA  
PT sequence encoding a heavy chain immunoglobulin linked to a peptide that  
PT targets a cellular compartment -  
XX  
PS Example 1; Fig 14; 81pp; English.  
XX  
CC The present invention relates to a method for modifying a plant to  
CC produce an antibody or an active fragment or derivative, or a protein  
CC functional equivalent, in a cellular compartment comprises introducing  
CC a DNA sequence encoding a heavy chain immunoglobulin, where the DNA is  
CC linked to promoters and provided with an additional sequence encoding a  
CC peptide capable of targeting heavy chain immunoglobulin to a cellular  
CC compartment. The method is used for producing a heavy chain  
CC immunoglobulin or an active fragment or derivative, or a protein that is  
CC functionally equivalent for increasing the pathogen resistance in a plant  
CC or to modulate metabolism in a plant. Under some circumstances it may be  
CC desirable to retain the antibody product with the plant rather than  
CC extracting and isolating the product. In particular, edible selected  
CC antigens may be used in a method of passively immunising an animal,  
CC preferably human, against the antigen, e.g., pathogenic organisms. The  
CC present DNA sequence encodes an anti-GUS antibody (denoted Clone18)

CC heavy chain variable domain (VH) attached to peptide linkers, myc and  
CC his6 tag. The GUS is beta-glucuronidase.

XX  
SQ Sequence 462 BP; 105 A; 136 C; 138 G; 83 T; 0 other;

Query Match 37.3%; Score 208; DB 22; Length 462;  
Best Local Similarity 72.8%; Pred. NO. 2.3e-47;  
Matches 303; Conservative 0; Mismatches 95; Indels 18; Gaps 2;

QY	12	caggtgcagctgcaggagtcagggggaggattggtgcaggctgggggctctctgagactc	71
Db	10	caggtgaaactgcagcagctctggggaggattggtgcaggctggggccctctgaggctc	69
QY	72	tctgtgcagcctcgggagcgcgccaccagtggtcactatggtatgggtggttc	131
Db	70	tctgtgcagcctctggcgcaccttca-----gtaactatgccgtgggtggttc	120
QY	132	cgccaggttccagggaaggagcgtgagttgtcgcagctattaggtggagtggtaaagag	191
Db	121	cgccaggtccagggaaggagcgtgagttgtcgtgctattagccgtgatggtgggcgc	180
QY	192	acatggtataaagactccgtgaaggccgattcaccatctccagagataacgccaaagact	251
Db	181	acatactatgcggactccgtgaaggccgattcgcgtctccagagactacgccgagaac	240
QY	252	acggtttatctgcaaatgaacagcctgaaacctgaagatacggccgtttatttgtgcc	311
Db	241	acggtgtatctgcaaatgaacagcctgaaacctgaggacacggccgtttattactgtaac	300
QY	312	gctcgaccggtccgcgtggtgatgatatccctgcgggttggtttgactactggggccag	371
Db	301	acaaggcctactggggccaggggaccaggtcacgt-----ctcctcagcgcac	351
QY	372	gggaccagggtcacccgtctcctcaggatctcatcaccatcaccatcacggatccac	427
Db	352	cacagcgaagacccccagctccgcggccgcccatcaccatcaccatcacggggccgc	407

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 13, 2002, 06:47:17 ; Search time 1846.07 Seconds  
(without alignments)  
6325.333 Million cell updates/sec

Title: us-09-742-690-1  
Perfect score: 558  
Sequence: 1 tcgagaaaagacaggtgcag.....ctcccagtgctgtaataag 558

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl:

- 1: gb\_ba:\*
- 2: gb\_htg:\*
- 3: gb\_in:\*
- 4: gb\_om:\*
- 5: gb\_ov:\*
- 6: gb\_pat:\*
- 7: gb\_ph:\*
- 8: gb\_pl:\*
- 9: gb\_pr:\*
- 10: gb\_ro:\*
- 11: gb\_sts:\*
- 12: gb\_sy:\*
- 13: gb\_un:\*
- 14: gb\_vi:\*
- 15: em\_ba:\*
- 16: em\_fun:\*
- 17: em\_hum:\*
- 18: em\_in:\*
- 19: em\_mu:\*
- 20: em\_om:\*
- 21: em\_or:\*
- 22: em\_ov:\*
- 23: em\_pat:\*
- 24: em\_ph:\*
- 25: em\_pl:\*
- 26: em\_ro:\*
- 27: em\_sts:\*
- 28: em\_un:\*
- 29: em\_vi:\*
- 30: em\_htg\_hum:\*
- 31: em\_htg\_inv:\*
- 32: em\_htg\_other:\*
- 33: em\_htgo\_inv:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
-----					

1	385.4	69.1	440	6	AX164056	AX164056 Sequence
2	385.4	69.1	888	6	AX180311	AX180311 Sequence
3	384.8	69.0	1154	6	AX164109	AX164109 Sequence
4	384.8	69.0	1172	6	AX164111	AX164111 Sequence
5	384	68.8	384	4	LGL236100	AJ236100 Lama glam
6	360	64.5	384	4	LGL236108	AJ236108 Lama glam
7	326.4	58.5	384	4	LGL236106	AJ236106 Lama glam
8	321.6	57.6	384	4	LGL236101	AJ236101 Lama glam
9	320	57.3	384	4	LGL236099	AJ236099 Lama glam
10	318.4	57.1	384	4	LGL236103	AJ236103 Lama glam
11	315.2	56.5	384	4	LGL236104	AJ236104 Lama glam
12	299.2	53.6	384	4	LGL236105	AJ236105 Lama glam
13	299.2	53.6	384	6	AX021690	AX021690 Sequence
14	291.2	52.2	714	6	AX021717	AX021717 Sequence
15	291.2	52.2	717	6	AX021721	AX021721 Sequence
16	291.2	52.2	750	6	AX021725	AX021725 Sequence
17	290.2	52.0	708	6	AX021727	AX021727 Sequence
18	254	45.5	378	4	AF442920	AF442920 Lama glam
19	236.6	42.4	375	4	AF441487	AF441487 Lama glam
20	234.8	42.1	378	4	AF442907	AF442907 Lama glam
21	234.8	42.1	378	4	AF442908	AF442908 Lama glam
22	233.4	41.8	372	4	AF442923	AF442923 Lama glam
23	233	41.8	387	4	LGL238059	AJ238059 Lama glam
24	232.8	41.7	375	6	A95262	A95262 Sequence 8
25	230.6	41.3	369	4	AF442927	AF442927 Lama glam
26	230.2	41.3	372	4	AF442915	AF442915 Lama glam
27	229.8	41.2	399	4	AF425245	AF425245 Lama glam
28	228.6	41.0	381	4	AF447920	AF447920 Lama glam
29	228.4	40.9	378	4	AF442917	AF442917 Lama glam
30	228.4	40.9	393	4	LGL238053	AJ238053 Lama glam
31	228	40.9	351	4	LGL429744	AJ429744 Lama glam
32	228	40.9	375	4	LGL236096	AJ236096 Lama glam
33	225.4	40.4	396	4	LGL238058	AJ238058 Lama glam
34	223.4	40.0	363	4	AF442916	AF442916 Lama glam
35	222.8	39.9	375	4	AF442912	AF442912 Lama glam
36	221.4	39.7	366	4	AF441486	AF441486 Lama glam
37	221.2	39.6	354	4	LGL237371	AJ237371 Lama glam
38	220.2	39.5	399	4	AF425248	AF425248 Lama glam
39	219.8	39.4	378	6	A95260	A95260 Sequence 6
40	217.2	38.9	399	4	LGL238054	AJ238054 Lama glam
41	216.4	38.8	354	4	LGL237328	AJ237328 Lama glam
42	216.2	38.7	369	4	AF442921	AF442921 Lama glam
43	215.8	38.7	381	4	AF442924	AF442924 Lama glam
44	215	38.5	366	4	LGL237388	AJ237388 Lama glam
45	214.2	38.4	360	4	AF447906	AF447906 Lama glam

ALIGNMENTS

RESULT	1					
AX164056						
LOCUS	AX164056	440 bp	DNA	linear	PAT 22-JUN-2001	
DEFINITION	Sequence 1 from Patent EP1118669.					
ACCESSION	AX164056					
VERSION	AX164056.1	GI:14545010				
KEYWORDS						
SOURCE	synthetic construct.					
ORGANISM	synthetic construct					
REFERENCE	1 (bases 1 to 440)					
AUTHORS	Frenken,L.G., van der Logt,C.P., Jobling,S.A. and Teh,Y.M.					
TITLE	Production of camelid antibodies in plants					
JOURNAL	Patent: EP 1118669-A 1 25-JUL-2001;					
FEATURES	UNILEVER PLC (GB) ; UNILEVER N.V. (NL)					
source	Location/Qualifiers					
	1..440					
	/organism="synthetic construct"					
	/db_xref="taxon:32630"					
	/note="VHH with peptide linker"					
	<1..420					
	/note="unnamed protein product"					
	/codon_start=1					
CDS						

/transl_table=11									
/protein_id="CAC42706.1"									
/db_xref="GI:14545011"									
/translation="QVQLQESGGGLVQAGGSLRLSCAASGRATSGHGHYGMGWFRRQVP									
GKEREFVAAIRWSGKETWYKDSVKGRFTISRDNAKTTVYLQMSLKPEDTAVYYCAAR									
PVRVDDISLPVGFYWGQGTQVTVSSQKLISEEDLN"									
BASE COUNT	103 a	103 c	135 g	99 t					
ORIGIN									
Query Match 69.1%; Score 385.4; DB 6; Length 440;									
Best Local Similarity 99.7%; Pred. No. 7.2e-111;									
Matches 386; Conservative 0; Mismatches 1; Indels 0; Gaps 0;									
QY	12	caggtgcagctgcaggagtcaggggagagattggtgcaggtcggggctctctgagactc	71						
Db	1	CAGGTGCAGCTGCAGGAGTCAGGGGAGGAGTTGGTCAGGCTGGGGCTCTCTGAGACTC	60						
QY	72	tcctgtcagcctcgggacgcgcaccagtggtcatggtcactatggtatgggtggttc	131						
Db	61	TCCTGTGCAGCCTCGGGACCGCCACCAGTGGTGCATGGTCACTATGGTATGGGCTGGTTC	120						
QY	132	cgccaggttcagggaagagcgtgagtttgtcgcagctattaggtggagtgtgtaaaagag	191						
Db	121	CGCCAGGTTCCAGGGAAGGAGCGCTGAGTTTGTTCGCAGCTATTAGGTGGAGTGGTAAAGAG	180						
QY	192	acatggtataaagactccgtgaaggccgattcaccatctccagagataaacgccaagact	251						
Db	181	ACATGGTATAAAGACTCCGTGAAGGGCCGATTTCACCATCTCCAGAGATAACGCCAAGACT	240						
QY	252	acggtttatctgcaaatgaacagcctgaaacctgaagatacagccgtttattattgtgcc	311						
Db	241	ACGTTTATCTGCAAAATGAACAGCCTGAAACCTGAAGATACGGCCGTTTATTATTGTGCC	300						
QY	312	gctcagccggtccgcgtggatgatatttccctgcggttgggttgactactggggccag	371						
Db	301	GCTCGACCGGTCCGCGTGGATGATATTTCCCTGCCGTTGGGTTTGACTACTGGGCCCAG	360						
QY	372	gggacccaggtcacccgtctcctcagga	398						
Db	361	GGGACCCAGGTCAACCGTCTCCTCAGAA	387						
RESULT 2									
AX180311									
LOCUS AX180311 888 bp DNA linear PAT 06-AUG-2001									
DEFINITION Sequence 9 from Patent WO0146364.									
ACCESSION AX180311									
VERSION AX180311.1 GI:15132274									
KEYWORDS synthetic construct.									
SOURCE synthetic construct.									
ORGANISM synthetic construct.									
REFERENCE 1 (bases 1 to 888)									
AUTHORS Howell,S., Little,J., van der Logt,C.P. and Parry,N.J.									
TITLE Method of delivering a benefit agent									
JOURNAL Patent: WO 0146364-A 9 28-JUN-2001;									
FEATURES Location/Qualifiers									
source									
1..888									
/organism="synthetic construct"									
/db_xref="taxon:32630"									
/note=" anti-RR6-VHH8-his-CBD"									
BASE COUNT	200 a	230 c	254 g	204 t					
ORIGIN									
Query Match 69.1%; Score 385.4; DB 6; Length 888;									
Best Local Similarity 96.1%; Pred. No. 7.7e-111;									
Matches 395; Conservative 0; Mismatches 16; Indels 0; Gaps 0;									
QY	12	caggtgcagctgcaggagtcaggggagagattggtgcaggtcggggctctctgagactc	71						
Db	1	CAGGTGCAGCTGCAGGAGTCAGGGGAGGAGTTGGTGCAGGCTGGGGCTCTCTGAGACTC	60						

QY	72	tcctgtcagcctcgggacgcgccaccagtggtcatggtcactatggtatgggtggttc	131																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																											
----	----	--	-----	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--



Db 67 CCTGTGCAGCCTCGGGACGGCCACCAGTGGTGCATGGTCACTATGGTATGGGCTGGTTCC 126

Qy 133 gccaggttccagggaagcgtgagtttgcgcagctattaggtggagtgtgtaaaagaga 192

Db 127 GCCAGGTTCCAGGGAAGGAGCGTGAGTTGTGCGAGCTATTAGGTGGAGTGGTAAAGAGA 186

Qy 193 catggtataaagactccgtgaagcgcgattcaccatctccagagataaacgccaagacta 252

Db 187 CATGGTATAAAGACTCCGTGAAGSGCCGATTACCATCTCCAGAGATAACGCCAAGACTA 246

Qy 253 catggtataaagactccgtgaagcgcgattcaccatctccagagataaacgccaagacta 252

Db 187 CATGGTATAAAGACTCCGTGAAGSGCCGATTACCATCTCCAGAGATAACGCCAAGACTA 246

Qy 253 cggtttatctgcaaatgaacagcctgaaacctgaagatacgcgcggtttattattgtgccg 312

Db 247 CGGTTTATCTGCAAAATGAACAGCCTGAAACCTGAAGATACGCGCGTTTATTATTGTGCCG 306

Qy 313 ctgcacgcgtccgcgtggatgatatttccctgccggttggtttgactactggtggccagg 372

Db 307 CTCGACCGTCCGCGTGGATGATATTCCCTGCCGTTGGGTTTGACTACTGGGGCCAGG 366

Qy 373 ggaccaggtcacccgtctcctcaggatc 400

Db 367 GGACCCAGGTCACCGTCTCCTCAGAAC 394

RESULT 4

AX164111

LOCUS AX164111 1172 bp DNA linear PAT 22-JUN-2001

DEFINITION Sequence 56 from Patent EP1118669.

ACCESSION AX164111

VERSION AX164111.1 GI:14545061

KEYWORDS synthetic construct.

SOURCE synthetic construct

ORGANISM artificial sequence.

REFERENCE 1 (bases 1 to 1172)

AUTHORS Frenken,L.G., van der Logt,C.P., Jobling,S.A. and Teh,Y.M.

TITLE Production of camelid antibodies in plants

JOURNAL Patent: EP 1118669-A 56 25-JUL-2001;

UNILEVER PLC (GB) ; UNILEVER N.V. (NL)

FEATURES

source

1. .1172

/organism="synthetic construct"

/db\_xref="taxon:32630"

/note="hinge-HCV33-CH2-CH3-SEKDEL"

3. .1154

/note="unnamed protein product"

/codon\_start=1

/transl\_table=11

/protein\_id="CAC42713.1"

/db\_xref="GI:14545062"

/translation="MEVQLQESGGGLVQAGGSLRLSCAASGRATSGHGHYGMGWFROV

PKEREFVAAIRWSKETWKDSVKGRFTISRDNKTTVYLOMNSLKPEDTAVYYCAA

RPVRVDDISLPVGFYWGQGTQVTVSSEPKTPKQPQPPQPNPTTESKQKCPAP

ELGGPSVFIFPPKPKDVLISISGRPEVTCVVDVQGEDPEVSFNWYIDGAEVRTANTR

PKEEFNSTYRVSVLPFIQHDLWLTGKEFKCKVNKALPAPIEKTISKAKGQTRPEQV

YALAPHREELAKDFVSVTCLVKGFYPPDINVEORNGQPESEGTATTPQLDNDGT

FLYSLKSVGKNTWQGETFTCVVMHEALHNHYTQKSITQSSGKSEKDEL"

BASE COUNT 300 a 343 c 326 g 203 t

ORIGIN

Query Match 69.0%; Score 384.8; DB 6; Length 1172;

Best Local Similarity 99.5%; Pred. No. 1.2e-110;

Matches 386; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 13 aggtgcagctgcaggagtcagggggagagattggtgcaggctgggggctctctgagactct 72

Db 7 AGGTGCAGCTGCAGGAGTCAGGGGAGGAGTGGTGCAGGCTGGGGGCTCTCTGAGACTCT 66

Qy 73 cctgtgcagctcgggacgcgccaccagtggtcatggtcactatggtatggctggttcc 132

Db 67 CCTGTGCAGCCTCGGGACGGCCACCAGTGGTCACTATGGTATGGGCTGGTTCC 126

Qy 133 gccaggttccagggaagcgtgagtttgcgcagctattaggtggagtgtgtaaaagaga 192

Db 127 GCCAGGTTCCAGGGAAGGAGCGTGAGTTGTGCGAGCTATTAGGTGGAGTGGTAAAGAGA 186

Qy 193 catggtataaagactccgtgaagcgcgattcaccatctccagagataaacgccaagacta 252

Db 187 CATGGTATAAAGACTCCGTGAAGSGCCGATTACCATCTCCAGAGATAACGCCAAGACTA 246

Qy 253 cggtttatctgcaaatgaacagcctgaaacctgaagatacgcgcggtttattattgtgccg 312

Db 247 CGGTTTATCTGCAAAATGAACAGCCTGAAACCTGAAGATACGCGCGTTTATTATTGTGCCG 306

Qy 313 ctgcacgcgtccgcgtggatgatatttccctgccggttggtttgactactggtggccagg 372

Db 307 CTCGACCGTCCGCGTGGATGATATTCCCTGCCGTTGGGTTTGACTACTGGGGCCAGG 366

Qy 373 ggaccaggtcacccgtctcctcaggatc 400

Db 367 GGACCCAGGTCACCGTCTCCTCAGAAC 394

RESULT 5

LGL236100

LOCUS LGL236100 384 bp DNA linear MAM 09-OCT-2000

DEFINITION Lama glama mRNA for Ig heavy chain variable region (r2).

ACCESSION AJ236100

VERSION AJ236100.1 GI:4165531

KEYWORDS Immunoglobulin heavy chain; immunoglobulin superfamily; variable region.

SOURCE llama.

ORGANISM Lama glama

REFERENCE 1 (bases 1 to 384)

AUTHORS van der Linden,R.H., de Geus,B., Frenken,G.J., Peters,H. and Verrips,C.T.

TITLE Improved production and function of llama heavy chain antibody fragments by molecular evolution

JOURNAL J. Biotechnol. 80 (3), 261-270 (2000)

MEDLINE 20403592

REFERENCE 2 (bases 1 to 384)

AUTHORS Frenken,L.G.J.

TITLE Direct Submission

JOURNAL Submitted (31-JUL-1998) Frenken L.G.J., Biotechnology, Unilever Research, Olivier van Noortlaan 120, 3133 AT Vlaardingen, THE NETHERLANDS

FEATURES

Location/Qualifiers

1. .384

/organism="Lama glama"

/db\_xref="taxon:9844"

/cell\_type="peripheral blood lymphocyte"

/rearranged

<1. .>384

/gene="r2"

1. .384

/gene="r2"

<1. .>384

/gene="r2"

/codon\_start=1

/protein\_id="CAAL5412.1"

/db\_xref="GI:4165532"

/translation="QVQLQESGGGLVQAGGSLRLSCAASGRATSGHGHYGMGWFROV

GKEREFVAAIRWSKETWKDSVKGRFTISRDNKTTVYLOMNSLKPEDTAVYYCAAR

PVRVDDISLPVGFYWGQGTQVTVSS"

BASE COUNT 80 a 93 c 124 g 87 t

ORIGIN

Query Match 68.8%; Score 384; DB 4; Length 384;

Best Local Similarity 100.0%; Pred. No. 2e-110;

Matches 384; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 12 caggtgcagctgcaggagtcagggggagagattggtgcaggctgggggctctctgagactc 71

Db 1 CAGGTGCAGCTGCAGGAGTCAGGGGGAGGATTGGTGCAGGCTGGGGCTCTCTGAGACTC 60

QY 72 tctgtgcagcctcggagcgcgccaccagtggtcatgtcactatggtatggctggttc 131  
|||||

Db 61 TCCTGTGCAGCCTCGGACCGCGCCACCAGTGGTCAATGCTCACTATGTTATGGGCTGGTTC 120

QY 132 cgccaggttccagggaagagcgtgagttgtgcagctattagtggtggtgtaaaagag 191  
|||||

Db 121 CGCCAGGTTCCAGGGAAGGAGCGTGAGTTTGTGCGCAGCTATTAGTGGAGTGSTAAAGAG 180

QY 192 acatggtataaagactccgtgaaggccgattccaccatctccagagataacgccaaagact 251  
|||||

Db 181 ACATGGTATAAGACTCCGTGAAGGGCCGATTCCACCATCTCCAGAGATAACGCCAAGACT 240

QY 252 acggtttatctgcaaatgaacagcctgaaacctgaagatacagccgctttattattgtgcc 311  
|||||

Db 241 ACGGTTTATCTGCAATGAACAGCCTGAAACCTGAAGATACGGCCGTTTATTATTGTGCC 300

QY 312 gctcgaccggtccgcgtgatgatatttccctgcgcggttggttgactactggggccag 371  
|||||

Db 301 GCTCGACCGGTCCGCGTGGATGATATTTCCTGCCGTTGGGTTTGACTACTGGGGCCAG 360

QY 372 gggaccaggtcacctcctctca 395  
|||||

Db 361 GGGACCCAGGTCACCGTCTCCTCA 384

RESULT 6

LGL236108

LOCUS LGL236108 384 bp DNA linear MAM 09-OCT-2000

DEFINITION Lama glama mRNA for Ig heavy chain variable region (r10).

ACCESSION AJ236108

VERSION AJ236108.1 GI:4165547

KEYWORDS immunoglobulin heavy chain; immunoglobulin superfamily; variable region.

SOURCE llama.

ORGANISM Lama glama

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Tylopoda; Camelidae; Lama.

REFERENCE 1 (bases 1 to 384)

AUTHORS van der Linden,R.H., de Geus,B., Frenken,G.J., Peters,H. and Verrips,C.T.

TITLE Improved production and function of llama heavy chain antibody fragments by molecular evolution

JOURNAL J. Biotechnol. 80 (3), 261-270 (2000)

MEDLINE 20403592

REFERENCE 2 (bases 1 to 384)

AUTHORS Frenken,L.G.J.

TITLE Direct Submission

JOURNAL Submitted (31-JUL-1998) Frenken L.G.J., Biotechnology, Unilever Research, Olivier van Noortlaan 120, 3133 AT Vlaardingen, THE NETHERLANDS

FEATURES

source 1..384

/organism="Lama glama"

/db\_xref="taxon:9844"

/cell\_type="peripheral blood lymphocyte"

/rearranged

<1..>384

/gene="r10"

/codon\_start=1

/protein\_id="CAA15420.1"

/db\_xref="GI:4165548"

/translation="QVQLQESGGGLVQAGGSLRLSCAASGRASSCHGGYGMGWFQRQVP GKREFVAAIRWSGKETWYVDSVKGRFTISRDNASTVYLQMSLKPEDTAVYYCAVR PVRVDISSPVGFDYWGQGTQTVVSS"

BASE COUNT 78 a 94 c 129 g 83 t

ORIGIN

Query Match 64.5%; Score 360; DB 4; Length 384;  
Best Local Similarity 96.1%; pred. No. 8e-103;  
Matches 369; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 12 caggtgcagctgcagagtcaggggagagattggtgcaggtggtcactatggtatggctgagactc 71  
|||||

Db 1 CAGGTGCAGCTGCAGAGTCAGGGGGAGGATTGGTGCAGGCTGGGGCTCTCTGAGACTC 60

QY 72 tctgtgcagcctcggagcgcgccaccagtggtcaggtggtcactatggtatggctggttc 131  
|||||

Db 61 TCCTGTGCAGCCTCGGACCGCGCCAGCAGTGGCCATGGTGGCTATGGCATGGGCTGGTTC 120

QY 132 cgccaggttccagggaagagcgtgagttgtgcagctattagtggtggtgtaaaagag 191  
|||||

Db 121 CGCCAGGTTCCAGGGAAGGAGCGTGAGTTTGTGCGCAGCTATTAGTGGAGCGSTAAAGAG 180

QY 192 acatggtataaagactccgtgaaggccgattccaccatctccagagataacgccaaagact 251  
|||||

Db 181 ACATGGTATGTAGACTCCGTGAAGGGCCGATTCCACCATCTCCAGAGACAACGCCAAGAGC 240

QY 252 acggtttatctgcaaatgaacagcctgaaacctgaagatacagccgctttattattgtgcc 311  
|||||

Db 241 ACGGTTTATCTGCAATGAACAGCCTGAAACCTGAAGATACGGCCGTTTATTATTGTGCC 300

QY 312 gctcgaccggtccgcgtgatgatatttccctgcgcggttggttgactactgagggccag 371  
|||||

Db 301 GTTCGACCGGTCCGCGTGGATGATATTTCACAGCCCGGTGGGTTTGACTACTGGGGCCAG 360

QY 372 gggaccaggtcacctcctctca 395  
|||||

Db 361 GGGACCCAGGTCACCGTCTCCTCA 384

RESULT 7

LGL236106

LOCUS LGL236106 384 bp DNA linear MAM 09-OCT-2000

DEFINITION Lama glama mRNA for Ig heavy chain variable region (r8).

ACCESSION AJ236106

VERSION AJ236106.1 GI:4165543

KEYWORDS immunoglobulin heavy chain; immunoglobulin superfamily; variable region.

SOURCE llama.

ORGANISM Lama glama

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Tylopoda; Camelidae; Lama.

REFERENCE 1 (bases 1 to 384)

AUTHORS van der Linden,R.H., de Geus,B., Frenken,G.J., Peters,H. and Verrips,C.T.

TITLE Improved production and function of llama heavy chain antibody fragments by molecular evolution

JOURNAL J. Biotechnol. 80 (3), 261-270 (2000)

MEDLINE 20403592

REFERENCE 2 (bases 1 to 384)

AUTHORS Frenken,L.G.J.

TITLE Direct Submission

JOURNAL Submitted (31-JUL-1998) Frenken L.G.J., Biotechnology, Unilever Research, Olivier van Noortlaan 120, 3133 AT Vlaardingen, THE NETHERLANDS

FEATURES

Location/Qualifiers

source 1..384

/organism="Lama glama"

/db\_xref="taxon:9844"

/cell\_type="peripheral blood lymphocyte"

/rearranged

<1..>384

/gene="r8"

1..384

/gene="r8"

<1..>384

/gene="r8"

/codon\_start=1

/protein\_id="CAA15418.1"

V\_region

gene

CDS

/db\_xref="GI:4165544"  
/translation="QVQLQSGGGLVQAGGSLRLSCVASGRFTSGHGGYGMGWFQRPV  
GKEREFVAIRWSGVTTYVDSVKGRFTISRDNKSTVYLQMSLKPEDTGVYYCAVR  
PVRVDISTPDPGFVYGQGTQVTVSS"  
BASE COUNT 76 a 94 c 126 g 88 t  
ORIGIN

Query Match 58.5%; Score 326.4; DB 4; Length 384;  
Best Local Similarity 90.6%; Pred. No. 3.5e-92;  
Matches 348; Conservative 0; Mismatches 36; Indels 0; Gaps 0;

QY 12 caggtgcagctgcaggagtcagggggaggattggtgcagctggggctctctgagactc 71  
|||||  
Db 1 CAGGTGCAGCTGCAGCAGTCAGGGGGAGGATTGTTGTCAGGCTGGGGGCTCTCTGAGACTC 60  
  
QY 72 tcctgtgcagcctcgggacgcccaccagtggtcatgtgcactatggtatgggctggttc 131  
|||||  
Db 61 TCTTGTGTAGCCTCGGGACGCACCTTCAGTGGCCATGGTGGCTATGGCATGGGCTGGTTC 120  
  
QY 132 cgccaggttccagggaagagcgtgagtttgcgcagctattaggtggagtggtaaaagag 191  
|||||  
Db 121 CGCCAGGTTCCAGGGAAGGCGGTGAGTTTGTCCGACGTATTAGGTGGAGCGGTGTACG 180  
  
QY 192 acatggtataaagactccgtgaaggccgattccaccatccagagataacgccaagact 251  
|||||  
Db 181 ACATATTATGTAGACTCCGTGAAGGGCCGATTACCATCTCCAGAGACAACGCCAAGAGC 240  
  
QY 252 acggtttatctgcaaatgaacagcctgaaacctgaagatacggcgtttattattgtgcc 311  
|||||  
Db 241 ACGGTGTATCTGCAANTGAACAGCCTGAAACCTGAAGATACGGGCGTTTATTATTGTGCA 300  
  
QY 312 gctgcaccggtccgcgtggatgatatattccctgcgcgttggtggttgactactggggccag 371  
|||||  
Db 301 GTTCGGCCGGTCCGCGTGGACGACATTTCACCCCGGATGGGTTTGTTTACTGGGGCCAG 360  
  
QY 372 gggacccaggtcacctctcctca 395  
|||||  
Db 361 GGGACCCAGGTCACCGTCTCCTCA 384

RESULT 8  
LGL236101  
LOCUS LGL236101 384 bp DNA linear MAM 10-FEB-1999  
DEFINITION Lama glama mRNA for Ig heavy chain variable region (r3).  
ACCESSION AJ236101  
VERSION AJ236101.1 GI:4165533  
KEYWORDS immunoglobulin heavy chain; immunoglobulin superfamily; variable region.  
SOURCE llama.  
ORGANISM Lama glama  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Tylopoda; Camelidae; Lama.  
REFERENCE 1 (bases 1 to 384)  
AUTHORS Frenken,L.G.J.  
TITLE Direct Submission  
JOURNAL Submitted (31-JUL-1998) Frenken L.G.J., Biotechnology, Unilever Research, Olivier van Noortlaan 120, 3133 AT Vlaardingen, THE NETHERLANDS  
FEATURES  
source Location/Qualifiers  
1. .384  
/organism="Lama glama"  
/db\_xref="taxon:9844"  
/cell\_type="peripheral blood lymphocyte"  
/rearranged  
<1. .>384  
/gene="r3"  
1. .384  
/gene="r3"  
<1. .>384  
/gene="r3"  
/codon\_start=1  
/protein\_id="CAA15413.1"  
  
V\_region  
gene  
CDS

/db\_xref="GI:4165534"  
/translation="QVQLQESGGGLVQAGDSLRLSCEASGRTSHGYGGMGWFQRPV  
GKERELVAAIRWSGLETYHKDSVKGRFTISRDNKTMVYLQMSLKPEDTAVYYCAAR  
TVRVVDISSPVGFAYWGQGTQVTVSS"  
BASE COUNT 82 a 97 c 121 g 84 t  
ORIGIN

Query Match 57.6%; Score 321.6; DB 4; Length 384;  
Best Local Similarity 89.8%; Pred. No. 1.2e-90;  
Matches 345; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

QY 12 caggtgcagctgcaggagtcagggggaggattggtgcagctggggctctctgagactc 71  
|||||  
Db 1 CAGGTGCAGCTGCAGGAGTCAGGGGGAGGATTGTTGTCAGGCTGGGGACTCTCTGAGACTC 60  
  
QY 72 tcctgtgcagcctcgggacgcccaccagtggtcatgtgcactatggtatgggctggttc 131  
|||||  
Db 61 TCCTGTGAAGCCTCGGGACGCACCTCCCATGGCTATGCTGGCTATGGCATGGGCTGGTTC 120  
  
QY 132 cgccaggttccagggaagagcgtgagtttgcgcagctattaggtggagtggtaaaagag 191  
|||||  
Db 121 CGCCAGGTTCCAGGGAAGGCGGTGAGCTTGTCCGACGTATTAGGTGGAGCGGTCTTGAA 180  
  
QY 192 acatggtataaagactccgtgaaggccgattccaccatccagagataacgccaagact 251  
|||||  
Db 181 ACATATCATAAAGACTCCGTGAAGGGCCGATTTCACCATCTCCAGAGACAACCAAGAAC 240  
  
QY 252 acggtttatctgcaaatgaacagcctgaaacctgaagatacggcgtttattattgtgcc 311  
|||||  
Db 241 ATGGTGTATCTACAAATGAACAGCCTGAAACCTGAGGACACGGCCGTTTATTATTGTGCA 300  
  
QY 312 gctgcaccggtccgcgtggatgatatattccctgcgcgttggtggttgactactggggccag 371  
|||||  
Db 301 GCTCGGACGGTTCGCGTGGTTGACATTTCAGCCCGTGGGTTTGCCTACTGGGGCCAG 360  
  
QY 372 gggacccaggtcacctctcctca 395  
|||||  
Db 361 GGGACCCAGGTCACCGTCTCCTCA 384

RESULT 9  
LGL236099  
LOCUS LGL236099 384 bp DNA linear MAM 10-FEB-1999  
DEFINITION Lama glama mRNA for Ig heavy chain variable region (r1).  
ACCESSION AJ236099  
VERSION AJ236099.1 GI:4165529  
KEYWORDS immunoglobulin heavy chain; immunoglobulin superfamily; variable region.  
SOURCE llama.  
ORGANISM Lama glama  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Tylopoda; Camelidae; Lama.  
REFERENCE 1 (bases 1 to 384)  
AUTHORS Frenken,L.G.J.  
TITLE Direct Submission  
JOURNAL Submitted (31-JUL-1998) Frenken L.G.J., Biotechnology, Unilever Research, Olivier van Noortlaan 120, 3133 AT Vlaardingen, THE NETHERLANDS  
FEATURES  
source Location/Qualifiers  
1. .384  
/organism="Lama glama"  
/db\_xref="taxon:9844"  
/cell\_type="peripheral blood lymphocyte"  
/rearranged  
<1. .>384  
/gene="r1"  
1. .384  
/gene="r1"  
<1. .>384  
/gene="r1"  
/codon\_start=1  
/protein\_id="CAA15411.1"  
  
V\_region  
gene  
CDS

/db\_xref="GI:4165530"  
/translation="QVQLQSGGGLVQAGDSLRSLSCASGRTSHGYGGMGWFQVP  
GKERELVAAIRWSGVETVHKDSVKGRFTISRDNKMMVYLQMSLKPEDTAVYYCAAR  
TVRVVDISSPVGFAYWGQGTQTVSS"  
BASE COUNT 81 a 98 c 122 g 83 t  
ORIGIN

Query Match 57.3%; Score 320; DB 4; Length 384;  
Best Local Similarity 89.6%; Pred. No. 3.8e-90;  
Matches 344; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

QY 12 caggtgcagctcaggaagtcagggggaggatcggtgcaggctgggggctctctgagactc 71  
|||||  
Db 1 CAGGTGCAGCTGCAGCAGTCAGGGGAGGATGGTGCAGGCTGGGGACTCTCTGAGACTC 60  
  
QY 72 tcctgtcagctcgggacgcgccaccagtggtcactatggtatgggctggttc 131  
|||||  
Db 61 TCCTGTGAAGCTCGGGACGCACCTCCCATGGCTATGGTGGCTATGGCATGGGCTGGTTC 120  
  
QY 132 cgccaggttcagggaaaggagcgtgagttgtcgcagctattaggtggagtggtaaaagag 191  
|||||  
Db 121 CGCCAGGTTCCAGGAAAGAGCGTGAGCTTTCGCAGCTATTAGGTGGAGCGGTGTGAA 180  
  
QY 192 acatggtataaagactccgtgaaggccgcatcaccatctccagagataacgccaaagact 251  
|||||  
Db 181 ACATATCATAAAGACTCCGTGAAGGCCGGTTCACCATCTCCAGAGACAACGCCAAGAAC 240  
  
QY 252 acggtttatctcaaatgaacagccctgaacctgaagatacgcgcgtttattattgtgcc 311  
|||||  
Db 241 ATGGTGTATCTACAAATGAACAGCCCTGAAACCTGAGGACACGGCCGTTTATTATTGTGCA 300  
  
QY 312 gctcgaccggtccgcgtggatgatatattccctgcgcggttggtggtttgactactcggggccag 371  
|||||  
Db 301 GCTCGGACGGTCCGCGTGGTTGACATTTCCAGCCCGGTTGGGTTTGCCTACTGGGGCCAG 360  
  
QY 372 gggaccaggtcacgctctcctca 395  
|||||  
Db 361 GGGACCCAGGTACCGTCTCCTCA 384

RESULT 10  
LGL236103  
LOCUS LGL236103 384 bp DNA linear MAM 10-FEB-1999  
DEFINITION Lama glama mRNA for Ig heavy chain variable region (r5).  
ACCESSION AJ236103  
VERSION AJ236103.1 GI:4165537  
KEYWORDS immunoglobulin heavy chain; immunoglobulin superfamily; variable region.  
SOURCE llama.  
ORGANISM Lama glama  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Tylopoda; Camelidae; Lama.  
REFERENCE 1 (bases 1 to 384)  
AUTHORS Frenken,L.G.J.  
TITLE Direct Submission  
JOURNAL Submitted (31-JUL-1998) Frenken L.G.J., Biotechnology, Unilever Research, Olivier van Noortlaan 120, 3133 AT Vlaardingen, THE NETHERLANDS

FEATURES  
Source  
1..384  
/organism="Lama glama"  
/db\_xref="taxon:9844"  
/cell\_type="peripheral blood lymphocyte"  
/rearranged  
<1..>384  
/gene="r5"  
1..384  
/gene="r5"  
<1..>384  
/gene="r5"  
/codon\_start=1  
/protein\_id="CAA15415.1"

V\_region  
gene  
CDS

/db\_xref="GI:4165538"  
/translation="QVQLQSGGGLVQAGSLRLSCASGRTSHGYGGMGWFQVP  
GKERELVAAIRWSGTSTVYADSVKGRFTISRDNKMMVYLQMSLKPEDTAVYHCAAR  
TVRVVDISSPVGFAYWGQGTQTVSS"  
BASE COUNT 77 a 100 c 125 g 82 t  
ORIGIN

Query Match 57.1%; Score 318.4; DB 4; Length 384;  
Best Local Similarity 89.3%; Pred. No. 1.2e-89;  
Matches 343; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

QY 12 caggtgcagctcaggaagtcagggggaggatcggtgcaggctgggggctctctgagactc 71  
|||||  
Db 1 CAGGTGCAGCTGCAGGAGTCAGGGGAGGATGGTGCAGGCTGGGGGCTCTCTGAGACTC 60  
  
QY 72 tcctgtcagctcgggacgcgccaccagtggtcactatggtatgggctggttc 131  
|||||  
Db 61 TCCTGTGAGCTCGGGACGCACCTTCCCATGGCTATGGTGGCTATGGCATGGGCTGGTTC 120  
  
QY 132 cgccaggttcagggaaaggagcgtgagttgtcgcagctattaggtggagtggtaaaagag 191  
|||||  
Db 121 CGCCAGGTTCCAGGAAAGAGCGTGAGCTTTCGCAGCAATTAGGTGGAGCGGTACTAGT 180  
  
QY 192 acatggtataaagactccgtgaaggccgattcaccatctccagagataaacgccaaagact 251  
|||||  
Db 181 ACATACTATGCAGACTCCGTGAAGGCCCGATTTCACCATCTCCAGAGACAACGTCAGAAGAC 240  
  
QY 252 acggtttatctcaaatgaacagccctgaacctgaagatacgcgcgtttattattgtgcc 311  
|||||  
Db 241 ATGGTGTATCTGCAAAATGAACAGCCTGAAACCTGAGGACACGGCCGTTTACCATTGTGCA 300  
  
QY 312 gctcgaccggtccgcgtggatgatatattccctgcgcggttggtggtttgactactcggggccag 371  
|||||  
Db 301 GCTCGGACGGTCCGCGTGGTTGACATTTCCAGCCCGGTTGGGTTTGCCTACTGGGGCCAG 360  
  
QY 372 gggaccaggtcacgctctcctca 395  
|||||  
Db 361 GGGACCCAGGTACCGTCTCCTCA 384

RESULT 11  
LGL236104  
LOCUS LGL236104 384 bp DNA linear MAM 10-FEB-1999  
DEFINITION Lama glama mRNA for Ig heavy chain variable region (r6).  
ACCESSION AJ236104  
VERSION AJ236104.1 GI:4165539  
KEYWORDS immunoglobulin heavy chain; immunoglobulin superfamily; variable region.  
SOURCE llama.  
ORGANISM Lama glama  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Tylopoda; Camelidae; Lama.  
REFERENCE 1 (bases 1 to 384)  
AUTHORS Frenken,L.G.J.  
TITLE Direct Submission  
JOURNAL Submitted (31-JUL-1998) Frenken L.G.J., Biotechnology, Unilever Research, Olivier van Noortlaan 120, 3133 AT Vlaardingen, THE NETHERLANDS

FEATURES  
Location/Qualifiers  
1..384  
/organism="Lama glama"  
/db\_xref="taxon:9844"  
/cell\_type="peripheral blood lymphocyte"  
/rearranged  
<1..>384  
/gene="r6"  
1..384  
/gene="r6"  
<1..>384  
/gene="r6"  
/codon\_start=1  
/protein\_id="CAA15416.1"

V\_region  
gene  
CDS



/db\_xref="GI:4165540"  
/translation="QVQLQQSGGGLVQTGGSLRLSCAASGRTSHGYGYGMGWFROVP  
GKERELVAAIRWSGRTYYADSVKGRFTISRDNVKNMVLYLQMSLKPEDTAVYYCAAR  
TVRVVDISSPVGFAYWGQGTQVTVSS"  
BASE COUNT 76 a 97 c 125 g 86 t  
ORIGIN

Query Match 56.5%; Score 315.2; DB 4; Length 384;  
Best Local Similarity 88.8%; Pred. No. 1.3e-88;  
Matches 341; Conservative 0; Mismatches 43; Indels 0; Gaps 0;  
  
QY 12 caggtgcagctgcaggagtcaggggaggattggtgcaggctggggctctctgagactc 71  
|||||  
Db 1 CAGGTGCAGCTGCAGCAGCTGCGGGGAGGATTGGTGCAGACTGGGGCTCTCTGAGACTC 60  
  
QY 72 tcctgtgcagcctcggaagcgccaccagtggtcatgtgcactatggtatgggctggttc 131  
|||||  
Db 61 TCCTGTGCAGCCTCGGGACCGCACTTCCCATGGGTATGGGTATGGCATGGGCTGGTTC 120  
  
QY 132 cgccaggttccagggaagagcgctgagtttgcgcagctattaggtggagtggttaaagag 191  
|||||  
Db 121 CGCCAGGTTCCAGGGAAGGAGCGTGAGCTTGTCGCAGCAATTAGGTGGAGCGGTATTAGT 180  
  
QY 192 acatggtataaagactccgtgaaggccgattcaccatctcagagataaacgccaaagact 251  
|||||  
Db 181 ACATACTATGCAGACTCCGTGAAGGGCCGATTACCATCTCCAGAGACAACGTCAGAAGAC 240  
  
QY 252 acggtttatctgcaaatgaacagcctgaaacctgaagatacggccgtttattattgtgcc 311  
|||||  
Db 241 ATGGTGTATCTGCAAAATGGACAGCCTGAAACCTGAGGACACGGCCGTTTATTATTGTGCA 300  
  
QY 312 gctcgaccggtccgcgtggatgatatattccctgccggttggtttgactactggtggccag 371  
|||||  
Db 301 GCTCGGACGGTCCGCGTGGTTGACATTTCACGCCCGGTTGGGTTTGCCCTACTGGGGCCAG 360  
  
QY 372 gggacccagggtcacccgtctcctca 395  
|||||  
Db 361 GGGACCCAGGTCACCGTCTCCTCA 384

RESULT 12  
LGL236105  
LOCUS LGL236105 384 bp DNA linear MAM 10-FEB-1999  
DEFINITION Lama glama mRNA for Ig heavy chain variable region (r7).  
ACCESSION AJ236105  
VERSION AJ236105.1 GI:4165541  
KEYWORDS immunoglobulin heavy chain; immunoglobulin superfamily; variable region.  
SOURCE llama.  
ORGANISM Lama glama  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Tylopoda; Camelidae; Lama.  
REFERENCE 1 (bases 1 to 384)  
AUTHORS Frenken,L.G.J.  
TITLE Direct Submission  
JOURNAL Submitted (31-JUL-1998) Frenken L.G.J., Biotechnology, Unilever Research, Olivier van Noortlaan 120, 3133 AT Vlaardingen, THE NETHERLANDS  
FEATURES  
source Location/Qualifiers  
1. .384  
/organism="Lama glama"  
/db\_xref="taxon:9844"  
/cell\_type="peripheral blood lymphocyte"  
/rearranged  
<1. .>384  
/gene="r7"  
1. .384  
/gene="r7"  
<1. .>384  
/gene="r7"  
/codon\_start=1  
/protein\_id="CAA15417.1"  
  
V\_region  
gene  
CDS

/db\_xref="GI:4165542"  
/translation="QVQLQESGGGLVQAGDSLRLSCAASGRTSHGYGYGMGWFROIP  
GKERELVAAIRWSGRNTYYADSVKGRFTISRDNVKDMLYLQMNSLKPEDTAVYTCAVR  
TVRVVDISSPVGFAYWGQGTQVTVSS"  
BASE COUNT 78 a 96 c 124 g 86 t  
ORIGIN

Query Match 53.6%; Score 299.2; DB 4; Length 384;  
Best Local Similarity 86.2%; Pred. No. 1.5e-83;  
Matches 331; Conservative 0; Mismatches 53; Indels 0; Gaps 0;  
  
QY 12 caggtgcagctgcaggagtcaggggaggattggtgcaggctggggctctctgagactc 71  
|||||  
Db 1 CAGGTGCAGCTGCAGGAGTCAGGGGAGGATTGGTGCAGCTGGGACTCTCTGAGACTC 60  
  
QY 72 tcctgtgcagcctcggaagcgccaccagtggtcatgtgcactatggtatgggctggttc 131  
|||||  
Db 61 TCCTGCGCGCCTCGGGACCGCACTTCTCATGGGTATGGTGGCTATGGCATGGGCTGGTTC 120  
  
QY 132 cgccaggttccagggaagagcgctgagtttgcgcagctattaggtggagtggttaaagag 191  
|||||  
Db 121 CGCCAAATTCCAGGGAAGGAGCGTGAGCTTGTCGCAGCAATTAGGTGGAGCGGTCCGTAAT 180  
  
QY 192 acatggtataaagactccgtgaaggccgattcaccatctcagagataaacgccaaagact 251  
|||||  
Db 181 ACATACTATGCAGACTCCGTGAAGGGCCGATTACCATCTCCAGAGACAACGTCAAGGAC 240  
  
QY 252 acggtttatctgcaaatgaacagcctgaaacctgaagatacggccgtttattattgtgcc 311  
|||||  
Db 241 ATGCTGTATCTGCAAAATGAACAGTTTGAACCTTGAGGACACGGCCGTTTACACTTGTGCA 300  
  
QY 312 gctcgaccggtccgcgtggatgatatattccctgccggttggtttgactactggtggccag 371  
|||||  
Db 301 GTTCGGACGGTCCGCGTGGTTGACATTTCACAGTCCGGTTGGGTTTGCCCTACTGGGGCCAG 360  
  
QY 372 gggacccagggtcacccgtctcctca 395  
|||||  
Db 361 GGGACCCAGGTCACCGTCTCCTCA 384  
  
RESULT 13  
AX021690  
LOCUS AX021690 384 bp DNA linear PAT 07-SEP-2000  
DEFINITION Sequence 21 from Patent WO9923221.  
ACCESSION AX021690  
VERSION AX021690.1 GI:10044968  
KEYWORDS llama.  
SOURCE Lama glama  
ORGANISM Lama glama  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Tylopoda; Camelidae; Lama.  
REFERENCE 1 (bases 1 to 384)  
AUTHORS Howell,S., Ledebouer,A.M., Frenken,L.G. and van der Logt,C.P.E.  
TITLE Multivalent antigen-binding proteins  
JOURNAL Patent: WO 9923221-A 21 14-MAY-1999;  
HOWELL STEVEN (GB); UNILEVER PLC (GB); LEDEBOER ADRIANUS MARINUS (NL); LOGT CORNELIS PAUL ERIK V D (NL); UNILEVER NV (NL); FRENKEN LEON GERARDUS JOSEPH (NL)  
FEATURES  
source Location/Qualifiers  
1. .384  
/organism="Lama glama"  
/db\_xref="taxon:9844"  
<1. .>384  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="CAC07784.1"  
/db\_xref="GI:10044969"  
/translation="QVQLQESGGGLVQAGDSLRLSCAASGRTSHGYGYGMGWFROIP  
GKERELVAAIRWSGRNTYYADSVKGRFTISRDNVKDMLYLQMNSLKPEDTAVYTCAVR  
TVRVVDISSPVGFAYWGQGTQVTVSS"  
BASE COUNT 78 a 96 c 124 g 86 t  
ORIGIN



Db 642 AGTTCGGACGGTCCGCGTGGTTGACATTTCCAGTCCGGTTGGGTTTGCCCTACTGGGGCCA 701

QY 371 ggggacccaggtcacc 386

||||||||||||||

Db 702 GGGGACCCAGGTCACC 717

Search completed: August 13, 2002, 12:30:12  
Job time: 20575 sec

OM of: US-09-742-690-1 to: Issued\_Patents\_AA:\* out\_format : pfs

Date: Aug 13, 2002 1:42 PM

About: Results were produced by the GenCore software, version 4.5,  
Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:

-MODEL=frame+n2p.model -DEV=xlp  
-Q=/cgn2\_1/USPTO\_spool/US09742690/runat\_12082002\_094914\_87/app\_query.fasta\_1.623  
-DB=Issued\_Patents\_AA -QFMT=fastan -SUFFIX=rai -GAPOP=12.000  
-GAPEXT=4.000 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000  
-QGAPOP=4.500 -QGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500  
-FGAPOP=6.000 -FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500  
-DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=blossum62  
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR\_SCORE=pct  
-THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs  
-NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09742690\_@CGN1\_1\_18 -NCPU=6 -ICPU=3 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -NO\_XLPXY -WAIT -THREADS=1

Search information block:

Query: US-09-742-690-1

Query length: 558

Database: Issued\_Patents\_AA:\*

Database sequences: 231628

Database length: 24425594

Search time (sec): 39.700000

score\_list:

Sequence	Strd Orig	Zscore	Escore	Len	! Documentation	
/cgn2_6/ptodata/2/iaa/6B_COMB.pep:US-08-983-607-28 +	403.50	706.85	5.3e-32	131	!	
/cgn2_6/ptodata/2/iaa/6A_COMB.pep:US-08-974-899-6 +	399.50	701.29	1.2e-31	113	!	S
/cgn2_6/ptodata/2/iaa/6B_COMB.pep:US-09-079-029-10 +	395.00	682.93	4.7e-31	312	!	
/cgn2_6/ptodata/2/iaa/5A_COMB.pep:US-08-129-930B-96 +	394.00	689.45	4.6e-31	139	!	
/cgn2_6/ptodata/2/iaa/6B_COMB.pep:US-08-134-346A-51 +	394.00	689.45	4.6e-31	139	!	
/cgn2_6/ptodata/2/iaa/6B_COMB.pep:US-08-976-288A-96 +	394.00	689.45	4.6e-31	139	!	
/cgn2_6/ptodata/2/iaa/5A_COMB.pep:US-07-988-925-11 +	393.50	690.15	4.9e-31	119	!	
/cgn2_6/ptodata/2/iaa/5B_COMB.pep:US-08-362-780-11 +	393.50	690.15	4.9e-31	119	!	
/cgn2_6/ptodata/2/iaa/5B_COMB.pep:US-07-934-373C-21 +	393.00	689.01	5.6e-31	122	!	
/cgn2_6/ptodata/2/iaa/6A_COMB.pep:US-08-437-642B-21 +	393.00	689.01	5.6e-31	122	!	
/cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:PCT-US93-07832-21 +	393.00	689.01	5.6e-31	122	!	
/cgn2_6/ptodata/2/iaa/6B_COMB.pep:US-09-025-769B-178 +	392.50	679.58	8.1e-31	281	!	
/cgn2_6/ptodata/2/iaa/5A_COMB.pep:US-08-331-398A-46 +	391.50	686.62	7.7e-31	119	!	
/cgn2_6/ptodata/2/iaa/5B_COMB.pep:US-08-331-397B-46 +	391.50	686.62	7.7e-31	119	!	
/cgn2_6/ptodata/2/iaa/5B_COMB.pep:US-08-759-804A-46 +	391.50	686.62	7.7e-31	119	!	
/cgn2_6/ptodata/2/iaa/6B_COMB.pep:US-09-227-693-46 +	391.50	686.62	7.7e-31	119	!	
/cgn2_6/ptodata/2/iaa/5B_COMB.pep:US-08-652-816A-14 +	391.00	685.99	8.6e-31	116	!	
/cgn2_6/ptodata/2/iaa/6B_COMB.pep:US-08-862-124-17 +	390.50	675.83	1.3e-30	287	!	
/cgn2_6/ptodata/2/iaa/6B_COMB.pep:US-08-862-124-14 +	390.50	675.24	1.3e-30	304	!	
/cgn2_6/ptodata/2/iaa/5B_COMB.pep:US-08-958-201-2 +	390.00	684.05	1.1e-30	118	!	S
/cgn2_6/ptodata/2/iaa/5B_COMB.pep:US-08-958-201-4 +	390.00	684.05	1.1e-30	118	!	S
/cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:PCT-US96-13152-4 +	389.00	668.72	2.1e-30	443	!	
/cgn2_6/ptodata/2/iaa/6B_COMB.pep:US-09-025-769B-38 +	388.00	680.34	1.7e-30	120	!	
/cgn2_6/ptodata/2/iaa/6B_COMB.pep:US-09-025-769B-63 +	388.00	680.34	1.7e-30	120	!	
/cgn2_6/ptodata/2/iaa/6B_COMB.pep:US-08-983-607-32 +	387.50	677.88	2.0e-30	140	!	
/cgn2_6/ptodata/2/iaa/6B_COMB.pep:US-08-983-607-36 +	387.00	678.92	2.1e-30	116	!	
/cgn2_6/ptodata/2/iaa/6B_COMB.pep:US-08-976-183A-55 +	386.50	675.40	2.6e-30	150	!	
/cgn2_6/ptodata/2/iaa/5B_COMB.pep:US-08-665-202-30 +	385.50	675.67	3.1e-30	123	!	
/cgn2_6/ptodata/2/iaa/5B_COMB.pep:US-08-428-197-1 +	385.50	675.50	3.1e-30	125	!	S
/cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:PCT-US93-10555-1 +	385.50	675.50	3.1e-30	125	!	
/cgn2_6/ptodata/2/iaa/6B_COMB.pep:US-08-579-378A-20 +	385.00	673.83	3.5e-30	135	!	
/cgn2_6/ptodata/2/iaa/5A_COMB.pep:US-08-398-612A-30 +	385.00	667.47	4.3e-30	251	!	
/cgn2_6/ptodata/2/iaa/5A_COMB.pep:US-08-398-611A-30 +	385.00	667.47	4.3e-30	251	!	
/cgn2_6/ptodata/2/iaa/5B_COMB.pep:US-08-491-334A-30 +	385.00	667.47	4.3e-30	251	!	
/cgn2_6/ptodata/2/iaa/6A_COMB.pep:US-09-027-449-27 +	385.00	667.47	4.3e-30	251	!	
/cgn2_6/ptodata/2/iaa/6A_COMB.pep:US-08-804-444A-27 +	385.00	667.47	4.3e-30	251	!	
/cgn2_6/ptodata/2/iaa/6B_COMB.pep:US-09-026-985-27 +	385.00	667.47	4.3e-30	251	!	
/cgn2_6/ptodata/2/iaa/6B_COMB.pep:US-08-398-613A-30 +	385.00	667.43	4.3e-30	252	!	
/cgn2_6/ptodata/2/iaa/5B_COMB.pep:US-08-653-402B-6 +	384.50	671.33	4.1e-30	158	!	
/cgn2_6/ptodata/2/iaa/5B_COMB.pep:US-08-653-402B-10 +	384.50	671.33	4.1e-30	158	!	

/cgn2\_6/ptodata/2/iaa/6B\_COMB.pep:US-08-983-607-46 + 383.50 672.64 4.7e-30 117  
/cgn2\_6/ptodata/2/iaa/5B\_COMB.pep:US-07-934-373C-4 + 383.00 671.50 5.3e-30 120  
/cgn2\_6/ptodata/2/iaa/6A\_COMB.pep:US-08-437-642B-4 + 383.00 671.50 5.3e-30 120  
/cgn2\_6/ptodata/2/iaa/6B\_COMB.pep:US-08-752-693A-4 + 382.50 670.87 5.9e-30 117  
/cgn2\_6/ptodata/2/iaa/5B\_COMB.pep:US-08-652-558-39 + 382.50 670.61 6.0e-30 120

seq\_name: /cgn2\_6/ptodata/2/iaa/6B\_COMB.pep:US-08-983-607-28

seq\_documentation\_block:

; Sequence 28, Application US/08983607

; Patent No. 6140470

; GENERAL INFORMATION:

; APPLICANT: Alan Garen

; APPLICANT: Xiaohong Cai

; TITLE OF INVENTION: Human Anti-Tumor Monoclonal Anti-

; NUMBER OF SEQUENCES: 51

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Department of Molecular Biophysics

; ADDRESSEE: and Biochemistry, Yale University

; STREET: 266 Whitney Avenue

; CITY: New Haven

; STATE: Connecticut

; COUNTRY: United States of America

; ZIP: 06520-8114

; COMPUTER READABLE FORM:

; MEDIUM TYPE: 3.5" 1.44 Mb diskette

; COMPUTER: IBM PC

; OPERATING SYSTEM: MS DOS

; SOFTWARE: Word Processing

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/983,607

; FILING DATE: April 27, 1998

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: PCT/IB96/01032

; FILING DATE: June 28, 1996

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Mary M. Krinsky

; REGISTRATION NUMBER: 32423

; REFERENCE/DOCKET NUMBER: OCR-679

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 203-773-9544

; TELEFAX: 203-773-1183

; INFORMATION FOR SEQ ID NO: 28:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 131 residues

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE:

; DESCRIPTION: polypeptide

; ORIGINAL SOURCE:

; ORGANISM: Homo sapiens (melanoma patient immu-

; INDIVIDUAL ISOLATE: peripheral blood lymphocytes

; IMMEDIATE SOURCE:

; LIBRARY: DM414 scFv antibodies obtained from

; LIBRARY: fUSE5 fusion phage construct

; CLONE: V13

; FEATURE:

; NAME/KEY: heavy chain

US-08-983-607-28

alignment\_scores:

Quality: 403.50 Length: 161  
Ratio: 3.736 Gaps: 4

Percent Similarity: 67.081 Percent Identity: 55.901

alignment\_block:

US-09-742-690-1 x US-08-983-607-28



Align seg 1/1 to: US-08-983-607-28 from: 1 to: 131

12 CAGGTGCAGCTGCAGGAGTCAGGGGGAGGATTGGTGCAGGCTGGGGGCTC 61  
:::||||| :::||||| :::||||| :::||||| :::|||||

1 GluValGlnLeuValGlnSerGlyGlyLeuValGlnProGlyGlySe 17

62 TCTGAGACTCTCCTGTGCAGCCTCGGGACGCCACCAGTGGTCAATGGTC 111  
||||| :::||||| :::||||| :::||||| :::|||||

17 rLeuArgLeuSerCysAlaAlaSerGlyPheThrPheSerSer..... 31

112 ACTATGGTATGGCTGGTTCGCCAGGTTCAGGGAGGAGCGTGAGTTT 161  
||||| :::||||| :::||||| :::||||| :::|||||

32 ..TyrAlaMetSerTrpValArgGlnAlaProGlyLysGlyLeuGluTrp 47

162 GTCGCAGCTATTAGGTGGTGGTAAAGAGACATGGTATAAAGACTCGCT 211  
||||| :::||||| :::||||| :::||||| :::|||||

48 ValSerAlaIleSerGlySerGlyGlySerThrTyrTyrAlaAspSerVa 64

212 GAAGGGCCGATTACCATCTCCAGAGATAACGCCAACACTACGGTTATC 261  
||||| :::||||| :::||||| :::||||| :::|||||

64 lLysGlyArgPheThrIleSerArgAspAsnSerLysAsnThrLeuTyrL 81

262 TGCRAATGAACAGCCTGAAACCTGAAGATACGGCCGTTTATTATTGTC 311  
||||| :::||||| :::||||| :::||||| :::|||||

81 euGlnMetAsnSerLeuArgAlaGluAspThrAlaValTyrTyrCysAla 97

312 .....GCTCGACCGGTCCCGTGGATGATATTTCCTGCCGGTTGGGTT 355  
::: |||

98 LysGlyValAlaPro.....Ph 103

356 TGACTACTGGGCCAGGGACCCAGGTCACCGTCTCCTCAGGATCTCATC 405  
||||| :::||||| :::||||| :::||||| :::|||||

103 eAspTyrTrpGlyGlnGlyThrProValThrValSerSerGlyGly.... 118

406 ACCATCACCATCAGCGATCCACCTCCATTGAAGGTCGTACCCAGTCTCAC 455

118 ..... 118

456 TACGGTCAGTGGTGGTATTGGTTACTCCGGT 488  
||| ||||| ||| :::||||

119 ...GlyGlySerGlyGlyGlyGlySerGlyGly 128

seq\_name: /cgn2\_6/ptodata/2/iaa/6A\_COMB.pep:US-08-974-899-6

seq\_documentation\_block:  
; Sequence 6, Application US/08974899  
; Patent No. 6037454  
; GENERAL INFORMATION:  
; APPLICANT: Presta, Leonard G.  
; APPLICANT: Jardieu, Paula M.  
; TITLE OF INVENTION: Humanized Anti-CD11a Antibodies  
; NUMBER OF SEQUENCES: 24  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Genentech, Inc.  
; STREET: 1 DNA Way  
; CITY: South San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94080  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: WinPatIn (Genentech)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/974,899  
; FILING DATE:  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 60/031971  
; FILING DATE: 11/27/96  
; ATTORNEY/AGENT INFORMATION:

NAME: Lee, Wendy M.  
REGISTRATION NUMBER: 40,378  
REFERENCE/DOCKET NUMBER: P1014R1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 650/225-1994  
TELEFAX: 650/952-9881  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 113 amino acids  
TYPE: Amino Acid  
TOPOLOGY: Linear  
US-08-974-899-6

alignment\_scores:  
Quality: 399.50 Length: 128  
Ratio: 4.035 Gaps: 2  
Percent Similarity: 77.344 Percent Identity: 64.844

alignment\_block:  
US-09-742-690-1 x US-08-974-899-6 ..  
Align seg 1/1 to: US-08-974-899-6 from: 1 to: 113

12 CAGGTGCAGCTGCAGGAGTCAGGGGGAGGATTGGTGCAGGCTGGGGGCTC 61  
:::||||| :::||||| :::||||| :::||||| :::|||||

1 GluValGlnLeuValGlnSerGlyGlyLeuValGlnProGlyGlySe 17

62 TCTGAGACTCTCCTGTGCAGCCTCGGGACGCCACCAGTGGTCAATGGTC 111  
||||| :::||||| :::||||| :::||||| :::|||||

17 rLeuArgLeuSerCysAlaAlaSerGlyPheThrPheSerSer..... 31

112 ACTATGGTATGGCTGGTTCGCCAGGTTCAGGGAGGAGCGTGAGTTT 161  
||||| :::||||| :::||||| :::||||| :::|||||

32 ..TyrAlaMetSerTrpValArgGlnAlaProGlyLysGlyLeuGluTrp 47

162 GTCGCAGCTATTAGGTGGTGGTAAAGAGACATGGTATAAAGACTCCGT 211  
||||| :::||||| :::||||| :::||||| :::|||||

48 ValSerValIleSerGlyAspGlyGlySerThrTyrTyrAlaAspSerVa 64

212 GAAGGGCCGATTACCATCTCCAGAGATAACGCCAACACTACGGTTATC 261  
||||| :::||||| :::||||| :::||||| :::|||||

81 euGlnMetAsnSerLeuArgAlaGluAspThrAlaValTyrTyrCysAla 97

312 GCTCGACCGGTCCCGTGGATGATATTTCCTGCCGGTTGGGTTGACTA 361  
||||| :::|||||

98 Arg.....GlyPheAspTy 102

362 CTGGGGCCAGGGACCCAGGTCACCGTCTCTCTCA 395  
||||| :::||||| :::||||| :::||||| :::|||||

102 rTrpGlyGlnGlyThrLeuValThrValSerSer 113

seq\_name: /cgn2\_6/ptodata/2/iaa/6B\_COMB.pep:US-09-079-029-10

seq\_documentation\_block:  
; Sequence 10, Application US/09079029  
; Patent No. 6342369  
; GENERAL INFORMATION:  
; APPLICANT: Adams, Camilia W.  
; APPLICANT: Ashkenazi, Avi J.  
; APPLICANT: Chuntharapai, Anan  
; APPLICANT: Kim, Kyung J.  
; TITLE OF INVENTION: Apo-2 Receptor  
; NUMBER OF SEQUENCES: 14  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Genentech, Inc.  
; STREET: 1 DNA Way  
; CITY: South San Francisco  
; STATE: California

```

; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/079,029
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Marschang, Diane L.
; REGISTRATION NUMBER: 35,600
; REFERENCE/DOCKET NUMBER: P1101R2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-5416
; TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 312 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
;
US-09-079-029-10

alignment_scores:
  Quality: 395.00      Length: 193
  Ratio: 3.211        Gaps: 7
  Percent Similarity: 63.731      Percent Identity: 51.295

alignment_block:
US-09-742-690-1 x US-09-079-029-10 ..

Align seg 1/1 to: US-09-079-029-10 from: 1 to: 312

15 GTGCAGCTGCAGGACTCAGGGGAGGATTGGTGCAGGCTGGGGGCTCTCT 64
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
41 ValGlnLeuValGluSerGlyGlyGlyLeuValGlnProGlyGlySerLe 57
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
65 GAGACTCTCTGTGCAGCCTCGGGACGGCGCCAGGAGTGGTCATGGTCACT 114
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
57 uArgLeuSerCysAlaAlaSerGlyPheThrPheSer.....T 71

115 ATGGTATGGCTGGTCCGCCAGGTTCCAGGAAGAGCGTGAGTTTGTC 164
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
71 yTrpMetSerTrpValArgGlnAlaProGlyLysGlyLeuGluTrpVal 87
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
165 GCAGCTATTAGTGGAGTGGTAAAGAGACATGGTATAAAGACTCCGTGAA 214
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
88 AlaAsnIleLysGlnAspGlySerGluLysTyrValAspSerVally 104
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
215 GGGCCGATTCCACCATCTCCAGAGATAACGCCACGACTACGGTTTATCTGC 264
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
104 sGlyArgPheThrIleSerArgAspAsnAlaLysAsnSerLeuTyrLeuG 121
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
265 AAATGAACAGCCTGAACCTGAAGATACGGCGGTTTATTATTGTCCGCT 314
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
121 InMetAsnSerLeuArgAlaGluAspThrAlaValTyrCysAlaArg 137
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
315 CGACCGGTCCGGTGGATGATATTTCCTGCGGTTGGGTTTGACTACTG 364
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
138 AspLeuLeuLysValLysGlySerSer..SerGlyTrp.PheAspProTr 153
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
365 GGGCCAGGGGACCCAGGTCACCGTCTCTCTCAGGATCTCATCACCATCAC 414
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
153 pGlyArgGlyThrThrValThrValSerSerGlyGly..... 165
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
415 ATCACGGATCCACCTCCATTGAAGGTGCTACCGAGTCTCACTACGGTCAG 464
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
166 .....GlyGly 167
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
465 TGTGGTGGTATTGGTTACTCCGGT..... 488
```

```

168 SerGlyGlyGlyGlySerGlyGlyGlyGlySerSerGluLeuThrGlnAs 184
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
489 .CCAACCGTCTGTGCCTCT.....GGTACCACCTGTCTCAG. 521
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
184 pProAlaValSerValAlaLeuGlyGlnThrValArgIleThrCysGlnG 201
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
522 .....GTTCTGAACCCCTTACTACTCC 542
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
201 lyAspSerLeuArgSerTyrTyrAla 209
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

seq_name: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:US-08-129-930B-96

seq_documentation_block:
; Sequence 96, Application US/08129930B
; Patent No. 5804187
; GENERAL INFORMATION:
; APPLICANT: do Couto Dr., Fernando J.R.
; APPLICANT: Ceriani Dr., Roberto L.
; APPLICANT: Peterson Dr., Jerry A.
; APPLICANT: Padlan Dr., Eduardo A.
; TITLE OF INVENTION: Analogue Peptides With Broad
; TITLE OF INVENTION: Carcinoma Specificity, and Kit and
; TITLE OF INVENTION: Diagnostic Vaccination and
; TITLE OF INVENTION: Therapeutic Methods
; NUMBER OF SEQUENCES: 96
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: V. AMZEL & ASSOC.
; STREET: 2055 NO. 5804187th Broadway, Suite 201
; CITY: Walnut Creek
; STATE: California
; COUNTRY: USA
; ZIP: 94596
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS 5.0
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/129,930B
; FILING DATE: September 30, 1993
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Amzel Ph.D., Viviana
; REGISTRATION NUMBER: 30,930
; REFERENCE/DOCKET NUMBER: CRFCC-008A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (510) 521-1333
; TELEFAX: (510) 521-3541
; TELEX: n.a.
; INFORMATION FOR SEQ ID NO: 96:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 139 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
;
US-08-129-930B-96

alignment_scores:
  Quality: 394.00      Length: 128
  Ratio: 3.863        Gaps: 2
  Percent Similarity: 79.688      Percent Identity: 63.281

alignment_block:
US-09-742-690-1 x US-08-129-930B-96 ..

Align seg 1/1 to: US-08-129-930B-96 from: 1 to: 139

12 CAGGTGCAGCTGCAGGAGTCAGGGGAGGATTGGTGCAGGCTGGGGCTC 61
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
20 GluValGlnMetValGluSerGlyGlyLeuValGlnProGlyGlySe 36
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
```

62 TCTGAGACTCTCTCTGTGCAGCCTCGGACCGCCGACGAGTGGTATGCTGTC 111  
|||||  
36 rLeuArgLeuSerCysAlaAlaSerGlyPheAlaPheSerSer..... 50  
112 ACTATGTTAGGCTGGTTCGCCAGGTTCCAGGGAAGGAGCGGTGAGTTT 161  
|||||  
51 ..TyrAlaMetSerTrpValArgGlnAlaProGlyLysGlyLeuGluTrp 66  
162 GTCGCAGCTATTAGTGGAGTGGTAAAGAGACATGGTATAAGACTCCGT 211  
|||||  
67 ValAlaGluIleSerSerGlyGlyAsnTyrAlaTyrTyrGlnAspThrVa 83  
212 GAAGGGCCGATTCACCATCTCCAGAGATAACGCCAAGACTACGGTTTATC 261  
|||||  
83 lThrGlyArgPheThrIleSerArgAspAsnSerLysAsnThrLeuTyrL 100  
262 TGCAATGAACAGCCTGAACCTGAAGATACGGCCGTTTATTATTGTGCC 311  
|||||  
100 euGlnMetAsnSerLeuArgAlaGluAspThrAlaValTyrTyrCysAla 116  
312 GCTCGACCGTCCGCGTGGATGATATTTCCTCCGCGTGGGTTGACTA 361  
|||||  
117 Arg.....GluAspTyrGlyIleProAlaTrpPheAlaTy 128  
362 CTGGGGCCAGGGACCCAGGTACCCGTCCTCTCA 395  
128 rTrpGlyGlnGlyThrLeuValThrValSerSer 139

seq\_name: /cgn2\_6/ptodata/2/iaa/6B\_COMB.pep:US-08-134-346A-51

seq\_documentation\_block:  
; Sequence 51, Application US/08134346A  
; Patent No. 6281335  
; GENERAL INFORMATION:  
; APPLICANT: do Couto, F.J.R.  
; APPLICANT: Ceriani, R.L.C.  
; APPLICANT: Petersen, J.A.  
; TITLE OF INVENTION: HYBRIDOMA AND ANTI-KC-4 HUMANIZED  
; TITLE OF INVENTION: MONOCLONAL ANTIBODY AND DNA AND RNA ENCODING IT, KIT AND  
; TITLE OF INVENTION: DIAGNOSTIC AND THERAPEUTIC METHODS  
; NUMBER OF SEQUENCES: 51  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Ostrager, Chong & Flaherty  
; STREET: 300 Park Avenue  
; CITY: New York  
; STATE: NY  
; COUNTRY: US  
; ZIP: 10022-7499  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette-3.50 inch, 1.44 Mb storage  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/134,346A  
; FILING DATE: 08-OCT-1993  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Onofrio, Dara L.  
; REGISTRATION NUMBER: 34,889  
; REFERENCE/DOCKET NUMBER: CLT 149,608  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-826-6565  
; TELEFAX: 212-826-5909  
; INFORMATION FOR SEQ ID NO: 51:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 139 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-134-346A-51

alignment\_scores:  
Quality: 394.00 Length: 128  
Ratio: 3.863 Gaps: 2  
Percent Similarity: 79.688 Percent Identity: 63.281  
alignment\_block:  
US-09-742-690-1 x US-08-134-346A-51 ..  
Align seg 1/1 to: US-08-134-346A-51 from: 1 to: 139  
12 CAGGTGCAGCTGCAGGAGTCCAGGGGAGGAGTGGTGCAGGCTGGGGGCTC 61  
:::|||||  
20 GluValGlnMetValGluSerGlyGlyLeuValGlnProGlyGlySe 36  
62 TCTGAGACTCTCTGTGCAGCCTCGGACCGCCGACGAGTGGTCTGCTGTC 111  
|||||  
36 rLeuArgLeuSerCysAlaAlaSerGlyPheAlaPheSerSer..... 50  
112 ACTATGTTAGTGGCTGGTTCGCCAGGTTCCAGGGAAGGAGCGGTGAGTTT 161  
|||||  
51 ..TyrAlaMetSerTrpValArgGlnAlaProGlyLysGlyLeuGluTrp 66  
162 GTCGCAGCTATTAGTGGAGTGGTAAAGAGACATGGTATAAGACTCCGT 211  
|||||  
67 ValAlaGluIleSerSerGlyGlyAsnTyrAlaTyrTyrGlnAspThrVa 83  
212 GAAGGGCCGATTCACCATCTCCAGAGATAACGCCAAGACTACGGTTTATC 261  
|||||  
83 lThrGlyArgPheThrIleSerArgAspAsnSerLysAsnThrLeuTyrL 100  
262 TGCAATGAACAGCCTGAACCTGAAGATACGGCCGTTTATTATTGTGCC 311  
|||||  
100 euGlnMetAsnSerLeuArgAlaGluAspThrAlaValTyrTyrCysAla 116  
312 GCTCGACCGTCCGCGTGGATGATATTTCCTCCGCGTGGGTTGACTA 361  
|||||  
117 Arg.....GluAspTyrGlyIleProAlaTrpPheAlaTy 128  
362 CTGGGGCCAGGGACCCAGGTACCCGTCCTCTCA 395  
128 rTrpGlyGlnGlyThrLeuValThrValSerSer 139

seq\_name: /cgn2\_6/ptodata/2/iaa/6B\_COMB.pep:US-08-976-288A-96

seq\_documentation\_block:  
; Sequence 96, Application US/08976288A  
; Patent No. 6315997  
; GENERAL INFORMATION:  
; APPLICANT: do Couto Dr., Fernando J.R.  
; APPLICANT: Ceriani Dr., Roberto L.  
; APPLICANT: Peterson Dr., Jerry A.  
; APPLICANT: Padlan Dr., Eduardo A.  
; TITLE OF INVENTION: Analogue Peptides With Broad  
; TITLE OF INVENTION: Carcinoma Specificity, and Kit and  
; TITLE OF INVENTION: Diagnostic Vaccination and  
; TITLE OF INVENTION: Therapeutic Methods  
; NUMBER OF SEQUENCES: 96  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pretty, Schroeder & Poplawski  
; STREET: 444 South Flower St., 19th Floor  
; CITY: Los Angeles  
; STATE: California  
; COUNTRY: USA  
; ZIP: 90071  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM-PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS 5.0  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/976,288A  
; FILING DATE: No. 6315997ember 21, 1997





262 TGCAAAATGAACAGCCCTGAAACCTGAAGATACGGCGGTTTATTATTGTGCC 311  
|||||  
81 euGlnMetAsnSerLeuArgAlaGluAspThrAlaValTyrTyrCysAla 97  
312 GCTCGACCGGTCCGGCTGGGATGATATTCCCTGCGGTTGGGTTTGACTA 361  
:::  
98 LysPheArg.....GlnTyrSerGlyGlyPheAspty 108  
362 CTGGGGCCAGGGACCCAGGTACCGTCTCCTCA 395  
|||||  
108 rTrpGlyGlnGlyThrLeuValThrValSerSer 119

seq\_name: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pep:US-08-362-780-11

seq\_documentation\_block:  
; Sequence 11, Application US/08362780  
; Patent No. 5968509  
; GENERAL INFORMATION:  
; APPLICANT: Gorman, Scott D  
; APPLICANT: Routledge, Edward G  
; APPLICANT: Waldmann, Herman  
; TITLE OF INVENTION: Antibody Preparation  
; NUMBER OF SEQUENCES: 26  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Nixon and Vanderhye pc  
; STREET: 8th Floor, 1100 No. 5968509th Glebe Road  
; CITY: Arlington  
; STATE: Virginia  
; COUNTRY: USA  
; ZIP: 22201  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/362,780  
; FILING DATE:  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/862,543  
; FILING DATE: 23-JUNE-1992  
; APPLICATION NUMBER: GB 9021679.7  
; FILING DATE: 05-OCT-1990  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: WO PCT/GB91/01726  
; FILING DATE: 04-OCT-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Mitchard, Leonard C  
; REGISTRATION NUMBER: 29009  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 7038164000  
; TELEFAX: 7038164100  
; INFORMATION FOR SEQ ID NO: 11:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 119 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-362-780-11

alignment\_scores:  
Quality: 393.50 Length: 128  
Ratio: 3.935 Gaps: 2  
Percent Similarity: 78.125 Percent Identity: 64.844

alignment\_block:  
US-09-742-690-1 x US-08-362-780-11 ..  
Align seg 1/1 to: US-08-362-780-11 from: 1 to: 119

12 CAGGTGCAGCTGCAGAGTCAAGGGGAGGATGTGTGAGCTGGGGGCTC 61  
:::  
1 GluValGlnLeuGluSerGlyGlyLeuValGlnProGlyGlyse 17  
62 TCTGAGACTCTCCTGTCAGCCTCGGGACGGCCACCAGTGGTGCATGGTC 111  
|||||  
17 rLeuArgLeuSerCysAlaAlaSerGlyPheThrPheSerSer..... 31  
112 ACTATGGTATGGGCTGGTTCGCCAGGTTCCAGGGAAGGAGCGTGAGTTT 161  
:::  
32 ..PheProMetAlaTrpValArgGlnAlaProGlyLysGlyLeuGluTrp 47  
162 GTCGCAGCTATTAGGTGGAGTGGTAAAGAGACATCGTATAAAGACTCCCGT 211  
|||||  
48 ValSerThrIleSerThrSerGlyGlyArgThrTyrTyrArgAspSerVa 64  
212 GAAGGGCCGATTCCACCATCTCCAGAGATAACGCCAAGACTACGGTTTATC 261  
|||||  
64 lLysGlyArgPheThrIleSerArgAspAsnSerLysAsnThrLeuTyrL 81  
262 TGCAAAATGAACAGCTGAAACCTGAAGATACGGCGGTTTATTATTGTGCC 311  
|||||  
81 euGlnMetAsnSerLeuArgAlaGluAspThrAlaValTyrTyrCysAla 97  
312 GCTCGACCGGTCCGGCTGGGATGATATTCCCTGCGGTTGGGTTTGACTA 361  
:::  
98 LysPheArg.....GlnTyrSerGlyGlyPheAspty 108  
362 CTGGGGCCAGGGACCCAGGTACCGTCTCCTCA 395  
|||||  
108 rTrpGlyGlnGlyThrLeuValThrValSerSer 119

seq\_name: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pep:US-07-934-373C-21

seq\_documentation\_block:  
; Sequence 21, Application US/07934373C  
; Patent No. 5821337  
; GENERAL INFORMATION:  
; APPLICANT: Paul J. Carter  
; APPLICANT: Leonard G. Presta  
; TITLE OF INVENTION: Immunoglobulin Variants  
; NUMBER OF SEQUENCES: 48  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Genentech, Inc.  
; STREET: 1 DNA Way  
; CITY: South San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94080  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: WinPatin (Genentech)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/934,373C  
; FILING DATE: 21-Aug-1992  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US92/05126  
; FILING DATE: 15-JUN-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/715272  
; FILING DATE: 14-JUN-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Lee, Wendy M.  
; REGISTRATION NUMBER: 40,378  
; REFERENCE/DOCKET NUMBER: P0709P2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 650/225-1994  
; TELEFAX: 650/952-9881  
; INFORMATION FOR SEQ ID NO: 21:

```
; SEQUENCE CHARACTERISTICS:
; LENGTH: 122 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
US-07-934-373C-21

alignment_scores:
  Quality: 393.00      Length: 128
  Ratio: 3.891        Gaps: 2
  Percent Similarity: 78.906      Percent Identity: 65.625

alignment_block:
US-09-742-690-1 x US-07-934-373C-21

Align seg 1/1 to: US-07-934-373C-21 from: 1 to: 122

12 CAGGTGCAGCTGCAGGAGTCAGGGGGAGGATTGGTGCAGGCTGGGGGCTC 61
:::|||||  |||||  |||||  |||||  |||||  |||||  |||||
1 GluValGlnLeuValGluSerGlyGlyGlyLeuValGlnProGlyGlySe 17

62 TCTGAGACTCTCCTGTGTCAGCTCGGGACGCGCCACCAAGCTGATGGTC 111
|||||  |||||  |||||  |||||  |||||  |||||  |||||
17 rLeuArgLeuSerCysAlaAlaSerGlyPheThrPheSerSer..... 31

112 ACTATGCTATGGCTGGTTCGCCAGGTTCCAGGGAAGGAGCGTGAGTTT 161
|||||  |||||  |||||  |||||  |||||  |||||  |||||
32 ..TyrAlaMetSerTrpValArgGlnAlaProGlyLysGlyLeuGluTrp 47

162 GTCCGACCTATTAGGTGGAGTGGTAAAGAGACATGGTATAAAGACTCCGT 211
|||||  |||||  |||||  |||||  |||||  |||||  |||||
48 ValSerValIleSerGlyAspGlySerThrTrpTyrAlaAspSerVa 64

212 GAAGGGCCGATTACCATCTCCAGAGATAACGCCAAGACTACGGTTTATC 261
|||||  |||||  |||||  |||||  |||||  |||||  |||||
64 lLysGlyArgPheThrIleSerArgAspAsnSerLysAsnThrLeuTyrL 81

262 TGCAATGAACAGCCTGAAACCTGAAGATACGGCCGTTTATTATTGTGCC 311
|||||  |||||  |||||  |||||  |||||  |||||  |||||
81 euGlnMetAsnSerLeuArgAlaGluAspThrAlaValTyrCysAla 97

312 GCTCGACCGTCCGCTGGTGATGATATTCCCTGCCGTTGGGTTTGACTA 361
|||  |||  |||||  |||||  |||||  |||||  |||||
98 ArgGlyArgVal.....GlyTyrSerLeuSerGlyLeuTyrAspTy 111

362 CTGGGGCCAGGGACCCAGGTACCGTCTCCTCA 395
|||||  |||||  |||||  |||||  |||||  |||||  |||||
111 rTrpGlyGlnGlyThrLeuValThrValSerSer 122

seq_name: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:US-08-437-642B-21

seq_documentation_block:
; Sequence 21, Application US/08437642B
; Patent No. 6054297
; GENERAL INFORMATION:
; APPLICANT: Paul J. Carter
; APPLICANT: Leonard G. Presta
; TITLE OF INVENTION: Immunoglobulin Variants
; NUMBER OF SEQUENCES: 47
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/437,642B
```

```
; FILING DATE: 09-May-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/934373
; FILING DATE: 21-AUG-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/146206
; FILING DATE: 17-NOV-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/05126
; FILING DATE: 15-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/715272
; FILING DATE: 14-JUN-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Lee, Wendy M.
; REGISTRATION NUMBER: 40,378
; REFERENCE/DOCKET NUMBER: P0709P2C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-1994
; TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 122 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
US-08-437-642B-21

alignment_scores:
  Quality: 393.00      Length: 128
  Ratio: 3.891        Gaps: 2
  Percent Similarity: 78.906      Percent Identity: 65.625

alignment_block:
US-09-742-690-1 x US-08-437-642B-21

Align seg 1/1 to: US-08-437-642B-21 from: 1 to: 122

12 CAGGTGCAGCTGCAGGAGTCAGGGGGAGGATTGGTGCAGGCTGGGGGCTC 61
:::|||||  |||||  |||||  |||||  |||||  |||||  |||||
1 GluValGlnLeuValGluSerGlyGlyGlyLeuValGlnProGlyGlySe 17

62 TCTGAGACTCTCCTGTGTCAGCTCGGGACGCGCCACCAAGCTGATGGTC 111
|||||  |||||  |||||  |||||  |||||  |||||  |||||
17 rLeuArgLeuSerCysAlaAlaSerGlyPheThrPheSerSer..... 31

112 ACTATGCTATGGCTGGTTCGCCAGGTTCCAGGGAAGGAGCGTGAGTTT 161
|||||  |||||  |||||  |||||  |||||  |||||  |||||
32 ..TyrAlaMetSerTrpValArgGlnAlaProGlyLysGlyLeuGluTrp 47

162 GTCCGACCTATTAGGTGGAGTGGTAAAGAGACATGGTATAAAGACTCCGT 211
|||||  |||||  |||||  |||||  |||||  |||||  |||||
48 ValSerValIleSerGlyAspGlySerThrTrpTyrAlaAspSerVa 64

212 GAAGGGCCGATTACCATCTCCAGAGATAACGCCAAGACTACGGTTTATC 261
|||||  |||||  |||||  |||||  |||||  |||||  |||||
64 lLysGlyArgPheThrIleSerArgAspAsnSerLysAsnThrLeuTyrL 81

262 TGCAATGAACAGCCTGAAACCTGAAGATACGGCCGTTTATTATTGTGCC 311
|||||  |||||  |||||  |||||  |||||  |||||  |||||
81 euGlnMetAsnSerLeuArgAlaGluAspThrAlaValTyrCysAla 97

312 GCTCGACCGTCCGCTGGTGATGATATTCCCTGCCGTTGGGTTTGACTA 361
|||  |||  |||||  |||||  |||||  |||||  |||||
98 ArgGlyArgVal.....GlyTyrSerLeuSerGlyLeuTyrAspTy 111

362 CTGGGGCCAGGGACCCAGGTACCGTCTCCTCA 395
|||||  |||||  |||||  |||||  |||||  |||||  |||||
111 rTrpGlyGlnGlyThrLeuValThrValSerSer 122

seq_name: /cgn2_6/ptodata/2/iaa/pctus_comb.pep:PCT-US93-07832-21
```

```
seq_documentation_block:
; Sequence 21, Application PC/TUS9307832
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; TITLE OF INVENTION: Immunoglobulin Variants
; NUMBER OF SEQUENCES: 40
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: patin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/07832
; FILING DATE: 19930820
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/715272
; FILING DATE: 14-JUN-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/05126
; FILING DATE: 15-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/934373
; FILING DATE: 21-AUG-1992
; ATTORNEY/AGENT INFORMATION:
; NAME:
; REGISTRATION NUMBER:
; REFERENCE/DOCKET NUMBER: 709P2PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE:
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 122 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
PCT-US93-07832-21
```

```
alignment_scores:
Quality: 393.00 Length: 128
Ratio: 3.891 Gaps: 2
Percent Similarity: 78.906 Percent Identity: 65.625
```

alignment\_block:

US-09-742-690-1 x PCT-US93-07832-21 ..

Align seg 1/1 to: PCT-US93-07832-21 from: 1 to: 122

```
12 CAGGTGCAGCTGCAGGAGTCAGGGGAGGATGGTGCAGGCTGGGGCTC 61
:::||||| ||||| ||||| ||||| ||||| ||||| |||||
1 GluValGlnLeuValGluSerGlyGlyLeuValGlnProGlyGlySe 17
62 TCTGAGACTCTCTGTGCAGCCTCGGGACGCCACCATGGTTCATGTC 111
||||| ||||| ||||| ||||| ||||| ||||| |||||
17 rLeuArgLeuSerCysAlaAlaSerGlyPheThrPheSerSer..... 31
112 ACTATGGTATGGCTGGTTCGCCAGGTTCCAGGGAAGAGCGTGAGTTT 161
||||| ||||| ||||| ||||| ||||| ||||| |||||
32 ..TyrAlaMetSerTrpValArgGlnAlaProGlyLysGlyLeuGluTrp 47
162 GTCGCAGCTATAGTGGAGTGGTAAAGACATGGTATAAGACTCCGT 211
||||| ||||| ||||| ||||| ||||| ||||| |||||
48 valSerValIleSerGlyAspGlyGlySerThrTyrTyrAlaAspSerVa 64
```

```
212 GAAGGGCCGATTCCACCATCTCCAGAGATAACGCCAAGACTACGGTTTATC 261
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
64 lLysGlyArgPheThrIleSerArgAspAsnSerLysAsnThrLeuTyrL 81
262 TGCAAAATGAACACCTGAAACCTGAAGATACGCCCGTTTATTATTGTGCC 311
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
81 euGlnMetAsnSerLeuArgAlaGluAspThrAlaValTyrTyrCysAla 97
312 GCTCGACCGGTCGGGTGGATGATATTTCTTCCTGCCGGTTGGGTTTGACTA 361
||| ||||| ||||| ||||| ||||| ||||| |||||
98 ArgGlyArgVal.....GlyTyrSerLeuSerGlyLeuTyrAspTyr 111
362 CTGGGGCCAGGGACCCAGGTCACCGTCTCCTCA 395
||||| ||||| ||||| ||||| ||||| ||||| |||||
111 rTrpGlyGlnGlyThrLeuValThrValSerSer 122
```

seq\_name: /cgn2\_6/ptodata/2/iaa/6B\_COMB.pcp:US-09-025-769B-178

```
seq_documentation_block:
; Sequence 178, Application US/09025769B
; Patent No. 6300064
; GENERAL INFORMATION:
; APPLICANT: Knappik, Achim
; APPLICANT: Pack, Peter
; APPLICANT: Ilag, Vic
; APPLICANT: Ge, Liming
; APPLICANT: Moroney, Simon
; APPLICANT: Plueckthun, Andreas
; TITLE OF INVENTION: Protein/(Poly)peptide libraries
; NUMBER OF SEQUENCES: 373
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: James F. Haley, Jr., Esq. c/o Fish & Neave
; STREET: 1251 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10021
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/025,769B
; FILING DATE: 18-FEB-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: EP 95 11 3021.0
; FILING DATE: 18-AUG-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: James F. Haley, Jr., Esq.
; REGISTRATION NUMBER: 27,794
; REFERENCE/DOCKET NUMBER: MORPHO/5
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)596-9000
; TELEFAX: (212)596-9090
; INFORMATION FOR SEQ ID NO: 178:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 281 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-025-769B-178
```

```
alignment_scores:
Quality: 392.50 Length: 164
Ratio: 3.634 Gaps: 4
Percent Similarity: 65.854 Percent Identity: 55.488
```

alignment\_block:

US-09-742-690-1 x US-09-025-769B-178 ..

Align seg 1/1 to: US-09-025-769B-178 from: 1 to: 281

6 AAAAGACAGGTGCAGCTGCAGGAGTCAGGGGAGGAGTGTGCAGGCTGG 55  
111 : : : : :  
24 LysAspGluValGlnLeuValGluSerGlyGlyGlyLeuValGlnProG1 40  
56 GGGCTCTCTGAGACTCTCCTGTGCAGCCTCGGGACGCGCCACAGTGGTC 105  
111 : : : : :  
40 yGlySerLeuArgLeuSerCysAlaAlaSerGlyPheThrPheSerSer. 56  
106 ATGGTCACTATGGTATGGCTGGTTCGCCAGGTTCCAGGGAAGGAGCGT 155  
111 : : : : :  
57 .....TyrAlaMetSerTrpValArgGlnAlaProGlyLysGlyLeu 70  
156 GAGTTTGTGCGCAGCTATTAGGTGGAGTGGTAAAGAGACATCGTATAAGA 205  
111 : : : : :  
71 GluTrpValSerAlaIleSerGlySerGlyGlySerThrTyrTyrAlaAs 87  
206 CTCGGTGAAGGCGCGATTACCATCTCCAGAGATACCGCAAGACTACGG 255  
111 : : : : :  
87 pSerValLysGlyArgPheThrIleSerArgAspAsnSerLysAsnThrL 104  
256 TTTATCTGCAAAATGAACAGCCCTGAAACCTGAAGATACGGCGGTTTATTAT 305  
: : : : :  
104 euTyrLeuGlnMetAsnSerLeuArgAlaGluAspThrAlaValTyrTyr 120  
306 TGTGCGCGCTCGACCGGTCGCGTGGATGATATTTCCCTGCGGTTGGTT 355  
111 : : : : :  
121 CysAlaArgTrpGly.....GlyAspGlyPh 129  
356 T.....GACTACTGGGGCCAGGGGACCCAGGTCACCGTCTCTCTCAG 396  
111 : : : : :  
129 eTyrAlaMetAspTyrTrpGlyGlnGlyThrLeuValThrValSerSera 146  
397 GATCTCATCACCATCACCGATCACCGATCCACCTCCATTGAAGTCTGTACC 446  
: : : : :  
146 laGly..... 147  
447 CAGTCTCACTACGGTCAGTGTGGTGGTATTGGTTACTCCGGT 488  
111 : : : : :  
148 .....GlyGlySerGlyGlyGlySerGlyGly 157

seq\_name: /cgn2\_6/ptodata/2/iaa/5A\_COMB.pep:US-08-331-398A-46

seq\_documentation\_block:  
; Sequence 46, Application US/08331398A  
; Patent No. 5608039  
; GENERAL INFORMATION:  
; APPLICANT: Pastan, Ira  
; APPLICANT: Willingham, Mark  
; APPLICANT: Fitzgerald, David  
; APPLICANT: Brinkmann, Ulrich  
; APPLICANT: Pai, Lee  
; TITLE OF INVENTION: Single Chain B3 Antibody Fusion Proteins  
; TITLE OF INVENTION: and Their Uses (as amended)  
; NUMBER OF SEQUENCES: 68  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend and Crew  
; STREET: One Market Plaza, Steuart Street Plaza  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94105-1492  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/331,398A  
; FILING DATE: 28-OCT-1994  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/767,331

; FILING DATE: 30-SEP-1991  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/596,289  
; FILING DATE: 12-OCT-1990  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Hunter, Tom  
; REGISTRATION NUMBER: 38,498  
; REFERENCE/DOCKET NUMBER: 015280-126110US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 543-9600  
; TELEFAX: (415) 543-5043  
; INFORMATION FOR SEQ ID NO: 46:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 119 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; FEATURE:  
; NAME/KEY: Protein  
; LOCATION: 1..119  
; OTHER INFORMATION: /note= "Human fetal immunoglobulin  
; OTHER INFORMATION: 56P1/CL Variable Heavy chain (V-H)"  
US-08-331-398A-46

alignment\_scores:  
Quality: 391.50 Length: 128  
Ratio: 3.876 Gaps: 2  
Percent Similarity: 78.906 Percent Identity: 64.062  
alignment\_block:  
US-09-742-690-1 x US-08-331-398A-46 ..  
Align seg 1/1 to: US-08-331-398A-46 from: 1 to: 119

12 CAGGTGCAGCTGCAGGAGTCAGGGGGAGGATTGGTGCAGGCTGGGGGCTC 61  
111 : : : : :  
1 GlnValGluLeuValGluSerGlyGlyGlyValValGlnProGlyArgSe 17  
62 TCTGAGACTCTCCTGTGCAGCCTCGGGACGCGCCACCATGTCATGGTC 111  
111 : : : : :  
17 rLeuArgLeuSerCysAlaAlaSerGlyPheThrPheSerSer..... 31  
112 ACTATGGTATGGCTGGTTCGCCAGGTTCCAGGGAAGGAGCGTGAGTTT 161  
111 : : : : :  
32 ..TyrAlaMetHisTrpValArgGlnAlaProGlyLysGlyLeuGluTrp 47  
162 GTCGACAGCTATTAGGTGGAGTGGTAAAGAGACATGGTATAAAGACTCCGT 211  
111 : : : : :  
48 ValAlaValIleSerTyrAspGlySerAsnLysTyrTyrAlaAspSerVa 64  
212 GAAGGCGCGATTCAACCATCTCCAGAGATAACGCCAAGACTACGGTTTATC 261  
111 : : : : :  
64 lLysGlyArgPheThrIleSerArgAspAsnSerLysAsnThrLeuTyrL 81  
262 TGCAAATGAACAGCCTGAACCTGAAGATACGGCCGTTTATTATTGTGCC 311  
111 : : : : :  
81 euGlnMetAsnSerLeuArgAlaGluAspThrAlaValTyrTyrCysAla 97  
312 GCTCGACCGTCCGCGTGGATGATATTTCCTCGCGGTTGGGTTTGACTA 361  
111 : : : : :  
98 ArgArgSerAlaArgThr.....TyrTyrPheAspTy 108  
362 CTGGGCGCAGGGACCCAGGTCACCGTCTCTCTCA 395  
111 : : : : :  
108 rTrpGlyGlnGlyThrLeuValThrValSerSer 119

seq\_name: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pep:US-08-331-397B-46

seq\_documentation\_block:  
; Sequence 46, Application US/08331397B  
; Patent No. 5981726



;; GENERAL INFORMATION:  
;; APPLICANT: Pastan, Ira  
;; APPLICANT: Benhar, Itai  
;; TITLE OF INVENTION: Chimeric and Mutationally Stabilized Tumor-  
;; TITLE OF INVENTION: Specific Antibody Fragments, Fusion Proteins, and Uses  
;; TITLE OF INVENTION: Thereof  
;; NUMBER OF SEQUENCES: 68  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Townsend and Townsend and Crew  
;; STREET: One Market Plaza, Steuart Street Plaza  
;; CITY: San Francisco  
;; STATE: California  
;; COUNTRY: USA  
;; ZIP: 94105-1492  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version #1.30  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/331,397B  
;; FILING DATE: 28-OCT-1994  
;; CLASSIFICATION: 435  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 07/767,331  
;; FILING DATE: 30-SEP-1991  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 07/596,289  
;; FILING DATE: 12-OCT-1990  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Hunter, Tom  
;; REGISTRATION NUMBER: 38,498  
;; REFERENCE/DOCKET NUMBER: 015280-126120US  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (415) 543-9600  
;; TELEFAX: (415) 543-5043  
;; INFORMATION FOR SEQ ID NO: 46:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 119 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS:  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: protein  
;; FEATURE:  
;; NAME/KEY: Protein  
;; LOCATION: 1..119  
;; OTHER INFORMATION: /note= "Human fetal immunoglobulin  
;; OTHER INFORMATION: 56p1'CL Variable Heavy chain (V-H)"  
US-08-331-397B-46

alignment\_scores:  
Quality: 391.50 Length: 128  
Ratio: 3.876 Gaps: 2  
Percent Similarity: 78.906 Percent Identity: 64.062

alignment\_block:  
US-09-742-690-1 x US-08-331-397B-46 ..

Align seg 1/1 to: US-08-331-397B-46 from: 1 to: 119

12 CAGGTGCAGCTGCAGAGTCCAGGGGAGGAGTGTGTCAGGCTGGGGCTC 61  
|||||:||||| |||||:|||||:||||| ||| ||  
1 GlnValGluLeuValGluSerGlyGlyValGlnProGlyArgse 17  
62 TCTGAGACTCTCTGTCAGCTCGGGACGCCAGTGGTCATGGTC 111  
|||||:|||||:|||||:|||||: |||: |||:  
17 rLeuArgLeuSerCysAlaAlaSerGlyPheThrPheSer..... 31  
112 ACTATGGTATGGCTGCTCCGCCAGGTTCCAGGAGGAGCGTGAGTTT 161  
|||||:||||| |||||:|||||:||||| |||:  
32 ..TyrAlaMethIstTrpValArgGlnAlaProGlyLysGlyLeuGluTrp 47

162 GTCCGACGCTATTAGTGGAGTGGTAAAGACACATGGTATAAGACTCCGT 211  
|||||:||||| :|||:|||||: ||| |||||  
48 ValAlaValIleSerTyrAspGlySerAsnLysTyrTyrAlaAspSerVa 64  
212 GAAGGGCCGATTCACCATCTCCAGAGATACGCCAAGACTACGGTTTATC 261  
|||||:|||||:|||||:|||||: |||: |||:  
64 lLysGlyArgPheThrIleSerArgAspAsnSerLysAsnThrLeuTyrL 81  
262 TGCAATGAACAGCCTGAACCTGAAGATACGCCGCTTATTATTGTGCC 311  
|||||:|||||: |||: |||||: |||||: |||||  
81 euGlnMetAsnSerLeuArgAlaGluAspThrAlaValTyrTyrCysAla 97  
312 GCTCGACCGTCCGCTGGATGATATTTCCCTGCCGTTGGTGTGACTA 361  
||| :|||: |||  
98 ArgArgSerAlaArgThr.....TyrTyrPheAspTyr 108  
362 CTGGGGCCAGGGACCCAGGTCACCGTCTCCTCA 395  
|||||:||||| |||||:|||||  
108 rTrpGlyGlnGlyThrLeuValThrValSerSer 119  
seq\_name: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pep:US-08-759-804A-46  
seq\_documentation\_block:  
; Sequence 46, Application US/08759804A  
; Patent No. 5990296  
; GENERAL INFORMATION:  
; APPLICANT: Pastan, Ira  
; APPLICANT: Willingham, Mark  
; APPLICANT: Fitzgerald, David J.  
; APPLICANT: Brinkmann, Ulrich  
; APPLICANT: Pai, Lee  
; TITLE OF INVENTION: Tumor-Specific Antibody Fragments,  
; TITLE OF INVENTION: Fusion Proteins, and Uses Thereof  
; NUMBER OF SEQUENCES: 68  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend and Crew LLP  
; STREET: Two Embarcadero Center, Eighth Floor  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94111-3834  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/759,804A  
; FILING DATE: 03-DEC-1996  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/331,398  
; FILING DATE: 28-OCT-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/767,331  
; FILING DATE: 30-SEP-1991  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/596,289  
; FILING DATE: 12-OCT-1990  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Weber, Ellen L.  
; REGISTRATION NUMBER: 32,762  
; REFERENCE/DOCKET NUMBER: 015280-126140US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 576-0200  
; TELEFAX: (415) 576-0300  
; INFORMATION FOR SEQ ID NO: 46:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 119 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein



OM of: US-09-742-690-1 to: A\_Geneseq\_032802:\* out\_format : pfs  
Date: Aug 13, 2002 1:41 PM  
About: Results were produced by the GenCore software, version 4.5,  
Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:  
-MODEL=frame+n2p.model -DEV=xlp  
-Q-/cgn2\_1/USPTO\_spool/US09742690/runat\_12082002\_094914\_76/app\_query.fasta\_1.623  
-DB-A\_Geneseq\_032802 -QFMT=fastan -SUFFIX=rag -GAPOP=12.000  
-GAPEXT=4.000 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000  
-QGAPOP=4.500 -QGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500  
-FGAPOP=6.000 -FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500  
-DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=blsum62  
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR\_SCORE=pct  
-THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs  
-NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US09742690@cgn1\_1\_87 -NCPU=6 -ICPU=3 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -NO\_XLPXY -WAIT -THREADS=1

Search information block:  
Query: US-09-742-690-1  
Query length: 558  
Database: A\_Geneseq\_032802:\*  
Database sequences: 747574  
Database length: 111073796  
Search time (sec): 85.160000

score_list:	Strd Orig	zScore	EScore	Len	! Documentation
Sequence					
/SIDS5/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:AAV79102 +			682.00	1163.22	9.5e-57
/SIDS5/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:AAE05281 +			682.00	1162.42	9.7e-57
/SIDS5/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:AAE05286 +			679.00	1147.69	2.4e-56
/SIDS5/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:AAE05287 +			679.00	1147.54	2.4e-56
/SIDS5/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:AAV79103 +			570.00	971.58	4.5e-46
/SIDS5/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:AAV15398 +			553.00	942.49	1.9e-44
/SIDS5/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:AAV15413 +			550.50	931.77	3.8e-44
/SIDS5/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:AAV79104 +			544.00	927.09	1.3e-43
/SIDS5/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:AAV15406 +			541.00	915.99	3.0e-43
/SIDS5/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:AAV15408 +			541.00	915.95	3.0e-43
/SIDS5/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:AAV15414 +			539.00	912.65	4.7e-43
/SIDS5/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:AAV08986 +			517.00	879.56	5.2e-41
/SIDS5/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:AAV08989 +			517.00	869.71	6.6e-41
/SIDS5/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:AAV08987 +			508.50	865.08	3.4e-40
/SIDS5/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:AAV08990 +			508.50	855.19	4.3e-40
/SIDS5/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:AAE05284 +			473.00	804.08	8.2e-37
/SIDS5/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:AAE05285 +			473.00	803.89	8.2e-37
/SIDS5/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:AAE05288 +			471.00	800.72	1.3e-36
/SIDS5/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:AAV39817 +			468.50	798.13	2.1e-36
/SIDS5/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:AAV15401 +			449.00	765.32	1.5e-34
/SIDS5/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:AAV08984 +			447.00	761.27	2.4e-34
/SIDS5/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:AAV28617 +			446.50	760.80	2.6e-34
/SIDS5/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:AAE05283 +			443.50	753.48	5.3e-34
/SIDS5/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:AAE07141 +			443.00	754.27	5.7e-34
/SIDS5/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:AAV15410 +			442.50	753.72	6.3e-34
/SIDS5/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:AAE05284 +			441.00	744.80	1.0e-33
/SIDS5/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:AAE05285 +			441.00	744.80	1.0e-33
/SIDS5/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:AAV28612 +			440.50	750.86	9.6e-34
/SIDS5/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:AAV15400 +			439.50	749.15	1.2e-33
/SIDS5/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:AAV70992 +			439.50	749.15	1.2e-33
/SIDS5/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:AAV08985 +			438.50	746.80	1.5e-33
/SIDS5/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:AAV28616 +			438.00	746.02	1.7e-33
/SIDS5/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:AAE05282 +			433.00	735.45	5.3e-33
/SIDS5/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:AAE10560 +			429.50	731.02	1.1e-32
/SIDS5/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:AAE07654 +			429.00	717.21	1.7e-32
/SIDS5/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:AAV41168 +			426.50	721.37	2.4e-32
/SIDS5/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:AAE10567 +			425.50	724.18	2.7e-32
/SIDS5/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:AAE10551 +			424.00	721.69	3.7e-32
/SIDS5/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:AAV28611 +			422.00	719.63	5.5e-32
/SIDS5/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:AAE07140 +			421.00	717.41	7.0e-32

/SIDS5/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:AAV28615 + 421.00 716.94 7.1e-32  
/SIDS5/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:AAE05281 + 421.00 716.63 7.1e-32  
/SIDS5/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:AAE10561 + 421.00 716.56 7.1e-32  
/SIDS5/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:AAV41173 + 420.00 710.39 9.9e-32  
/SIDS5/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:AAE072413 + 420.00 701.81 1.2e-31

seq\_name: /SIDS5/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:AAV79102

seq\_documentation\_block:

ID AAV79102 standard; Protein; 128 AA.

AC AAV79102;

DT 23-MAY-2000 (first entry)

XX Llama heavy chain variable region recombinant fragment R2.

XX Antibody; heavy chain variable region; protein secretion;

KW glycosylation; llama; Saccharomyces cerevisiae.

OS Lama guanicoe glama.

XX WO200005389-A2.

PD 03-FEB-2000.

XX 08-JUL-1999; 99WO-EP04919.

XX 20-JUL-1998; 98EP-0202432.

PA (UNIL ) UNILEVER NV.

PA (UNIL ) UNILEVER PLC.

XX Frenken LGJ, Sagt C, Verkleij AJ, Verrips CT;

PI WPI; 2000-182710/16.

XX High level recombinant production of heterologous protein with increased secretory efficiency in lower eukaryotic cells, useful for producing lysozyme, cutinase, or antibodies -

XX Example 3d; Page 37; 55pp; English.

XX The present sequence is that of llama anti-Azo dye RR6 heavy chain variable region (VH) recombinant protein R2. Anti-RR6 VH DNA was amplified by PCR (see AAZ58489-91), ligated into Yeast multicopy integration vectors and expressed in Saccharomyces cerevisiae. Anti-RR6 VH fragments R2, R5 and R7 were isolated and sequenced (see AAY79102-04). Secretion of the fragments was improved by the introduction at specific sites of N-glycosylation sites. The invention generally relates to the high level recombinant production of proteins, such as antibody fragments, in lower eukaryotic host (yeast or mould) cells. At least one N-glycosylation site is introduced into the protein at a location between the N-terminus and first hydrophobic region. The site is selected to increase the level of secretion without affecting the protein's functionality.

XX Sequence 128 AA;

alignment\_scores:

Quality: 682.00 Length: 128  
Ratio: 5.328 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-742-690-1 x AAV79102 ..

Align seg 1/1 to: AAV79102 from: 1 to: 128

12 CAGGTGCAGCTGCAGGAGTCAGGGGGAGGATTTGGTCAGGCTGGGGCTC 61

|||||  
1 GlnValGlnLeuGlnGluSerGlyGlyLeuValGlnAlaGlyGlyse 17

62 TCTGAGACTCTCCTGTGCAGCCCTCGGAGCGGCCACCAGTGGTCAATGGTC 111  
17 rLeuArgLeuSerCysAlaAlaSerGlyArgAlaThrSerGlyHisGlyH 34  
112 ACTATGGTATGGGTGGTCCGCCAGGTTCAGGGGAAGGAGCGTGAGTTT 161  
34 iSTyrGlyMetGlyTrpPheArgGlnValProGlyLysGluArgGluPhe 50  
162 GTCGACGCTATTAGGTGGAGTGGTAAAGAGACATGGTATAAGACTCCGT 211  
51 ValAlaAlaIleArgTrpSerGlyLysGluThrTrpTyrLysAspSerVa 67  
212 GAAGGGCCGATTACCATCTCCAGAGATAACGCCAAGACTACGGTTTATC 261  
67 lLysGlyArgPheThrIleSerArgAspAsnAlaLysThrValTyrL 84  
262 TGCAAAATGAACAGCCTGAAACCTGAAGATACGGCCGTTTATTATTGTGC 311  
84 euGlnMetAsnSerLeuLysProGluAspThrAlaValTyrCysAla 100  
312 GCTCGACCGGTCCCGGTGGATGATATTTCCCTGCCGGTGGGTTTGACTA 361  
101 AlaArgProValArgValAspAspIleSerLeuProValGlyPheAspTy 117  
362 CTGGGGCCAGGGACCCAGGTACCGTCTCCTCA 395  
117 rTrpGlyGlnGlyThrGlnValThrValSerSer 128  
seq\_name: /SIDS5/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:AAE05281  
seq\_documentation\_block:  
ID AAE05281 standard; Protein; 139 AA.  
XX  
AC AAE05281;  
XX  
DT 18-SEP-2001 (first entry)  
XX  
DE Llama HCV33 antibody heavy chain variable domain with myc tag.  
XX  
KW Camelid; llama; heavy chain immunoglobulin; pathogen resistance;  
KW metabolism modulator; passive immunisation; heavy chain variable domain;  
KW VH; anti-RR6 antibody; HCV33; azo-dye.  
XX  
OS Chimeric - Lama guanicoe glama.  
OS Chimeric - Unidentified.  
XX  
PN EP1118669-A2.  
XX  
PD 25-JUL-2001.  
XX  
PF 08-DEC-2000; 2000EP-0310997.  
XX  
PR 17-DEC-1999; 99EP-0310188.  
XX  
PA (UNIL ) UNILEVER PLC.  
PA (UNIL ) UNILEVER NV.  
XX  
PI Frenken LGJ, Van Der Logt CPE, Jobling SA, Teh Y;  
XX  
DR WPI; 2001-427157/46.  
DR N-PSDB; AAD10052.  
XX  
PT Modifying a plant to produce an antibody useful for increasing pathogen  
PT resistance or to modulate metabolism comprises introducing a DNA  
PT sequence encoding a heavy chain immunoglobulin linked to a peptide that  
PT targets a cellular compartment -  
XX  
PS Example 1; Fig 1; 8lpp; English.  
XX  
CC The present invention relates to a method for modifying a plant to  
CC produce an antibody or an active fragment or derivative, or a protein  
CC functional equivalent, in a cellular compartment comprises introducing

CC a DNA sequence encoding a heavy chain immunoglobulin, where the DNA is  
CC linked to promoters and provided with an additional sequence encoding a  
CC peptide capable of targeting heavy chain immunoglobulin to a cellular  
CC compartment. The method is used for producing a heavy chain  
CC immunoglobulin or an active fragment or derivative, or a protein that is  
CC functionally equivalent for increasing the pathogen resistance in a plant  
CC or to modulate metabolism in a plant. Under some circumstances it may be  
CC desirable to retain the antibody product with the plant rather than  
CC extracting and isolating the product. In particular, edible selected  
CC antigens may be used in a method of passively immunising an animal,  
CC preferably human, against the antigen, e.g., pathogenic organisms. The  
CC present sequence is llama (camelid) anti-RR6 (RR6 is an azo-dye)  
CC antibody (designated as HCV33) heavy chain variable domain (VH) attached  
CC with an peptide linker, myc peptide.  
XX  
SQ Sequence 139 AA;  
alignment\_scores:  
Quality: 682.00 Length: 128  
Ratio: 5.328 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000  
alignment\_block:  
US-09-742-690-1 x AAE05281 ..  
Align seg 1/1 to: AAE05281 from: 1 to: 139  
12 CAGGTGCAGCTGCAGGAGTCAAGGGGAGGATTGGTCAGGCTGGGGGCTC 61  
1 GlnValGlnLeuGlnGluSerGlyGlyGlyLeuValGlnAlaGlyGlyse 17  
62 TCTGAGACTCTCCTGTGCAGCCTCGGGACGGCCACCAGTGGTCAATGGTC 111  
17 rLeuArgLeuSerCysAlaAlaSerGlyArgAlaThrSerGlyHisGlyH 34  
112 ACTATGGTATGGGTGGTCCGCCAGGTTCAGGGGAAGGAGCGTGAGTTT 161  
34 iSTyrGlyMetGlyTrpPheArgGlnValProGlyLysGluArgGluPhe 50  
162 GTCGACGCTATTAGGTGGAGTGGTAAAGAGACATGGTATAAGACTCCGT 211  
51 ValAlaAlaIleArgTrpSerGlyLysGluThrTrpTyrLysAspSerVa 67  
212 GAAGGGCCGATTACCATCTCCAGAGATAACGCCAAGACTACGGTTTATC 261  
67 lLysGlyArgPheThrIleSerArgAspAsnAlaLysThrValTyrL 84  
262 TGCAAAATGAACAGCCTGAAACCTGAAGATACGGCCGTTTATTATTGTGC 311  
84 euGlnMetAsnSerLeuLysProGluAspThrAlaValTyrCysAla 100  
312 GCTCGACCGGTCCCGGTGGATGATATTTCCCTGCCGGTGGGTTTGACTA 361  
101 AlaArgProValArgValAspAspIleSerLeuProValGlyPheAspTy 117  
362 CTGGGGCCAGGGACCCAGGTACCGTCTCCTCA 395  
117 rTrpGlyGlnGlyThrGlnValThrValSerSer 128  
seq\_name: /SIDS5/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:AAE05286  
seq\_documentation\_block:  
ID AAE05286 standard; Protein; 377 AA.  
XX  
AC AAE05286;  
XX  
DT 18-SEP-2001 (first entry)  
XX  
DE Llama HCV33 VH region with heavy chain constant regions.  
XX  
KW Camelid; llama; heavy chain immunoglobulin; pathogen resistance;  
KW metabolism modulator; passive immunisation; heavy chain variable domain;



```
KW VH; anti-RR6 antibody; HCV33; azo-dye.
XX
OS Chimeric - Lama guanicoe glama.
OS Chimeric - Unidentified.
XX
FH Key Location/Qualifiers
FT Region 1..129
FT /note= "Corresponds to Llama HVC33 VH region"
FT Region 130..158
FT /note= "Corresponds to hinge region"
FT Region 159..268
FT /note= "Corresponds to CH2 heavy chain constant region"
FT Region 269..377
FT /note= "Corresponds to CH3 heavy chain constant region"
XX
PN EP1118669-A2.
XX
PD 25-JUL-2001.
XX
PF 08-DEC-2000; 2000EP-0310997.
XX
PR 17-DEC-1999; 99EP-0310188.
XX (UNIL ) UNILEVER PLC.
PA (UNIL ) UNILEVER NV.
XX
PI Frenken LGJ, Van Der Logt CPE, Jobling SA, Teh Y;
XX
DR WPI; 2001-427157/46.
DR N-PSDB; AAD10057.
XX
PT Modifying a plant to produce an antibody useful for increasing pathogen
PT resistance or to modulate metabolism comprises introducing a DNA
PT sequence encoding a heavy chain immunoglobulin linked to a peptide that
PT targets a cellular compartment -
XX
PS Example 12; Fig 18; 81pp; English.
XX
CC The present invention relates to a method for modifying a plant to
CC produce an antibody or an active fragment or derivative, or a protein
CC functional equivalent, in a cellular compartment comprising introducing
CC a DNA sequence encoding a heavy chain immunoglobulin, where the DNA is
CC linked to promoters and provided with an additional sequence encoding a
CC peptide capable of targeting heavy chain immunoglobulin to a cellular
CC compartment. The method is used for producing a heavy chain
CC immunoglobulin or an active fragment or derivative, or a protein that is
CC functionally equivalent for increasing the pathogen resistance in a plant
CC or to modulate metabolism in a plant. Under some circumstances it may be
CC desirable to retain the antibody product with the plant rather than
CC extracting and isolating the product. In particular, edible selected
CC antigens may be used in a method of passively immunising an animal,
CC preferably human, against the antigen, e.g., pathogenic organisms. The
CC present sequence is a llama (camelid) anti-RR6 (RR6 is an azo-dye)
CC antibody (designated as HCV33) heavy chain variable domain (VH)
CC linked to the non-classical heavy chain constant regions denoted as
CC hinge-CH2-CH3.
XX
SQ Sequence 377 AA;
```

```
alignment_scores:
  Quality: 679.00 Length: 128
  Ratio: 5.305 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 99.219
```

```
alignment_block:
US-09-742-690-1 x AAE05286 ..
```

Align seg 1/1 to: AAE05286 from: 1 to: 377

```
12 CAGGTGCAGCTGCAGGAGTCAGGGGAGGATTGGTCAGGCTGGGGGCTC 61
:::|||||
2 GluValGlnLeuGlnGluSerGlyGlyLeuValGlnAlaGlyGlyse 18
```

```
62 TCTGAGACTCTCTGTGCAGCCTCGGACGCGCCACCGAGTGCATGGTC 111
|||||
18 rLeuArgLeuSerCysAlaAlaSerGlyArgAlaThrSerGlyHisGlyH 35
|||||
112 ACTATGGTATGGGCTGGTTCGCGCAGGTTCCAGGGAAGGAGCGTGAGTTT 161
|||||
35 iTyrglyMetGlyTrpPheArgGlnValProGlyLysGluArgGluPhe 51
|||||
162 GTCGCAGCTATTAGTGGAGTGGTAAAGAGACATGGTATATAAGACTCCGT 211
|||||
52 ValAlaAlaIleArgTrpSerGlyLysGluThrTrpTyrLysaspSerVa 68
|||||
212 GAAGGGCGGATTCACCATCTCCAGAGATAACGCCAAGACTACGGTTTATC 261
|||||
68 lLysGlyArgPheThrIleSerArgAspAsnAlaLysThrValTyrL 85
|||||
262 TGCAAAATGAACAGCCTGAAACCTGAAGATACGGCGTTTATTATTGTGCC 311
|||||
85 euGlnMetAsnSerLeuLysProGluAspThrAlaValTyrCysAla 101
|||||
312 GCTCGACCGTCGCGTGGATGATATTCCTCGCGGTTGGGTTGACTA 361
|||||
102 AlaArgProValArgValAspAspIleSerLeuProValGlyPheAspty 118
|||||
362 CTGGGGCCAGGGACCCAGGTCAACCGTCTCCTCA 395
|||||
118 rTrpGlyGlnGlyThrGlnValThrValSerSer 129
```

seq\_name: /SIDS5/gcgdata/geneseq/geneseq-embl/AA2001.DAT:AAE05287

seq\_documentation\_block:

ID AAE05287 standard; Protein; 383 AA.

AC AAE05287;

XX

DT 18-SEP-2001 (first entry)

XX

DE Llama HCV33 VH region with ER retention signal and hinge-CH2-CH3.

XX

KW Camelid; llama; heavy chain immunoglobulin; pathogen resistance;  
KW metabolism modulator; passive immunisation; heavy chain variable domain;  
KW VH; anti-RR6 antibody; HCV33; azo-dye.

XX

OS Chimeric - Lama guanicoe glama.

OS Chimeric - Unidentified.

XX

FH Key Location/Qualifiers

FT Region 1..129

FT /note= "Corresponds to Llama HVC33 VH region"

FT Region 130..158

FT /note= "Corresponds to hinge region"

FT Region 159..268

FT /note= "Corresponds to CH2 heavy chain constant region"

FT Region 269..377

FT /note= "Corresponds to CH3 heavy chain constant region"

FT Region 378..383

FT /note= "Corresponds to ER retention signal sequence"

XX

PN EP1118669-A2.

XX

PD 25-JUL-2001.

XX

PF 08-DEC-2000; 2000EP-0310997.

XX

PR 17-DEC-1999; 99EP-0310188.

XX

PA (UNIL ) UNILEVER PLC.

PA (UNIL ) UNILEVER NV.

XX

PI Frenken LGJ, Van Der Logt CPE, Jobling SA, Teh Y;

XX

DR WPI; 2001-427157/46.

DR N-PSDB; AAD10058.  
XX  
PT Modifying a plant to produce an antibody useful for increasing pathogen  
PT resistance or to modulate metabolism comprises introducing a DNA  
PT sequence encoding a heavy chain immunoglobulin linked to a peptide that  
PT targets a cellular compartment -  
XX  
PS Example 12; Fig 19; 81pp; English.  
XX  
CC The present invention relates to a method for modifying a plant to  
CC produce an antibody or an active fragment or derivative, or a protein  
CC functional equivalent, in a cellular compartment comprising introducing  
CC a DNA sequence encoding a heavy chain immunoglobulin, where the DNA is  
CC linked to promoters and provided with an additional sequence encoding a  
CC peptide capable of targeting heavy chain immunoglobulin to a cellular  
CC compartment. The method is used for producing a heavy chain  
CC immunoglobulin or an active fragment or derivative, or a protein that is  
CC functionally equivalent for increasing the pathogen resistance in a plant  
CC or to modulate metabolism in a plant. Under some circumstances it may be  
CC desirable to retain the antibody product with the plant rather than  
CC extracting and isolating the product. In particular, edible selected  
CC antigens may be used in a method of passively immunising an animal,  
CC preferably human, against the antigen, e.g., pathogenic organisms. The  
CC present sequence is a llama (camelid) anti-RR6 (RR6 is an azo-dye)  
CC antibody (designated as HCV33) heavy chain variable domain (VH) linked  
CC to the non-classical heavy chain constant regions denoted as  
CC hinge-CH2-CH3 and endoplasmic reticulum (ER) retention signal.  
XX  
SQ Sequence 383 AA;

alignment\_scores:  
Quality: 679.00 Length: 128  
Ratio: 5.305 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 99.219  
alignment\_block:  
US-09-742-690-1 x AAE05287 ..  
Align seg 1/1 to: AAE05287 from: 1 to: 383  
12 CAGGTGCAGCTGCAGGAGTCAGGGGGAGGATTGGTGCAGGGCTGGGGGCTC 61  
:::|||||  
2 GluValGlnLeuGlnGluSerGlyGlyLeuValGlnAlaGlyGlySe 18  
62 TCTGAGACTCTCCTGTGCAGCCTCGGGAGCGGCCACCGAGTGGTCATGGTC 111  
|||||  
18 rLeuArgLeuSerCysAlaAlaSerGlyArgAlaThrSerGlyHisGlyH 35  
112 ACTATGGTATGGGCTGGTTCGCCAGGTTCCAGGGAAGGAGCGTGACTTT 161  
|||||  
35 iTyrglyMetGlyTrpPheArgGlnValProGlyLysGluArgGluPhe 51  
162 GTCGCAGCTATTAGGTGGAGTGGTAAAGACATGGTATAAAGACTCCGT 211  
|||||  
52 ValAlaAlaIleArgTrpSerGlyLysGluThrTrpTrpLysAspSerVa 68  
212 GAAGGGCCGATTCACCATCTCCAGAGATAACGCCAAGACTACGGTTTATC 261  
|||||  
68 lLysGlyArgPheThrIleSerArgAspAsnAlaLysThrThrValTyrl 85  
262 TGCAATGAACAGCCTGAAACCTGAAGATACGGCGTTTATTATTGTGCC 311  
|||||  
85 euGlnMetAsnSerLeuLysProGluAspThrAlaValTyrtyrCysAla 101  
312 GCTGACCGGTCGCGGTGGATGATATTTCCTCGCGGTTGGGTTGACTA 361  
|||||  
102 AlaArgProValArgValAspAspIleSerLeuProValGlyPheAspTy 118  
362 CTGGGGCCAGGGGACCCAGGTACCGTCTCCTCA 395  
|||||  
118 rTrpGlyGlnGlyThrGlnValThrValSerSer 129

seq\_name: /SIDS5/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:AA79103  
seq\_documentation\_block:  
ID AAY79103 standard; Protein; 128 AA.  
XX  
AC AAY79103;  
XX  
DT 23-MAY-2000 (first entry)  
XX  
DE Llama heavy chain variable region recombinant fragment R5.  
XX  
KW Antibody; heavy chain variable region; protein secretion;  
glycosylation; llama; Saccharomyces cerevisiae.  
XX  
OS Lama guanicoe glama.  
XX  
PN WO200005389-A2.  
XX  
PD 03-FEB-2000.  
XX  
PF 08-JUL-1999; 99WO-EP04919.  
XX  
PR 20-JUL-1998; 98EP-0202432.  
XX  
PA (UNIL ) UNILEVER NV.  
PA (UNIL ) UNILEVER PLC.  
XX  
PI Frenken LGJ, Sagt C, Verkleij AJ, Verrips CT;  
XX  
DR WPI; 2000-182710/116.  
XX  
PT High level recombinant production of heterologous protein with  
increased secretory efficiency in lower eukaryotic cells, useful for  
producing lysozyme, cutinase, or antibodies -  
XX  
PS Example 3d; Page 37; 55pp; English.  
XX  
CC The present sequence is that of llama anti-Azo dye RR6 heavy chain  
variable region (VH) recombinant protein R5. Anti-RR6 VH DNA was  
amplified by PCR (see AAZ58489-91), ligated into yeast multicopy  
integration vectors and expressed in Saccharomyces cerevisiae.  
CC Anti-RR6 VH fragments R2, R5 and R7 were isolated and sequenced  
(see AAY79102-04). Secretion of the fragments was improved by the  
introduction at specific sites of N-glycosylation sites. The  
invention generally relates to the high level recombinant production  
of proteins, such as antibody fragments, in lower eukaryotic host  
(yeast or mould) cells. At least one N-glycosylation site is  
introduced into the protein at a location between the N-terminus  
and first hydrophobic region. The site is selected to increase the  
level of secretion without affecting the protein's functionality.  
XX  
SQ Sequence 128 AA;

alignment\_scores:  
Quality: 570.00 Length: 128  
Ratio: 4.790 Gaps: 0  
Percent Similarity: 92.969 Percent Identity: 85.938  
alignment\_block:  
US-09-742-690-1 x AAY79103 ..  
Align seg 1/1 to: AAY79103 from: 1 to: 128  
12 CAGGTGCAGCTGCAGGAGTCAGGGGGAGGATTGGTGCAGGGCTGGGGGCTC 61  
|||||  
1 GlnValGlnLeuGlnGluSerGlyGlyLeuValGlnAlaGlyGlySe 17  
62 TCTGAGACTCTCCTGTGCAGCCTCGGGAGCGGCCACCGAGTGGTCATGGTC 111  
|||||  
17 rLeuArgLeuSerCysAlaAlaSerGlyArgThrSerHisGlyTyrgly 34  
112 ACTATGGTATGGGCTGGTTCGCCAGGTTCCAGGGAAGGAGCGTGACTTT 161

|||||  
34 lyTyrGlyMetGlyTrpPheArgGlnValProGlyLysGluArgGluLeu 50  
162 GTCGCAGCTATTAGGTGGAGTGGTAAAGAGACATGGTATAAAGACTCCTG 211  
|||||  
51 ValAlaAlaIleArgTrpSerGlyThrSerThrTyrTyrAlaAspSerVa 67  
212 GAAGGGCCGATTACCATCTCCAGAGATAACGCCAAGACTACGGTTTATC 261  
|||||  
67 lLysGlyArgPheThrIleSerArgAspAsnValLysAsnMetValTyrL 84  
262 TGCAAAATGAACAGCCTGAAACCTGAAGATACGGCGGTTTATTATTCGCC 311  
|||||  
84 euGlnMetAsnSerLeuLysProGluAspThrAlaValTyrHisCysAla 100  
312 GCTCGACCGTCCTCGGTGGATGATATTCCTGCGGTTGGGTTTACTA 361  
|||||  
101 AlaArgThrValArgValValAspIleSerSerProValGlyPheAlaTy 117  
362 CTGGGGCCAGGGACCCAGGTCAACCGTCTCCTCA 395  
|||||  
117 rTrpGlyGlnGlyThrGlnValThrValSerSer 128

seq\_name: /SIDS5/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:AA15398

seq\_documentation\_block:  
ID AAY15398 standard; Protein; 128 AA.

XX  
AC AAY15398;  
XX  
DT 21-JUL-1999 (first entry)  
XX  
DE Protein encoded by PstI/BstEII fragment of plasmid pUR4538.  
XX  
KW Multivalent antigen binding protein; single polypeptide chain;  
KW single domain binding site; diagnosis; therapy; targeting;  
KW immunoassay; cross-linking; agglutination; purification;  
KW phage inactivation; detergent.  
XX  
OS Synthetic.  
XX  
PN WO9923221-A2.  
XX  
PD 14-MAY-1999.  
XX  
PF 27-OCT-1998; 98WO-EP06991.  
XX  
PR 27-OCT-1997; 97EP-0308538.

XX  
PA (UNIL ) UNILEVER NV.  
PA (UNIL ) UNILEVER PLC.  
XX  
PI Frenken LGJ, Howell S, Ledebroer AM, Van Der Logt CPE;  
XX  
DR WPI; 1999-313342/26.  
DR N-PSDB; AAX59602.

XX  
PT Multivalent antigen binding protein useful for inactivation of  
(bacterio)phages or viruses

XX  
PS Disclosure; Fig 2; 89pp; English.  
XX  
CC The specification describes multivalent antigen binding proteins  
CC comprising a single polypeptide chain comprising, in series, two or  
CC more single domain binding sites. The multivalent antigen proteins  
CC are useful in applications where antibodies have been used in the  
CC prior art, including diagnosis, therapy, targeting, immunoassays,  
CC cross-linking methods including agglutination, or for purification  
CC processes. It is also useful for inactivation of (bacterio)phages  
CC or viruses, and in detergents. The present sequence represents the  
CC heavy chain variable domain of an anti-RR6 antibody from a llama.  
XX  
SQ Sequence 128 AA;

alignment\_scores:  
Quality: 553.00 Length: 128  
Ratio: 4.726 Gaps: 0  
Percent Similarity: 91.406 Percent Identity: 82.812  
alignment\_block:  
US-09-742-690-1 x AAY15398 ..  
Align seg 1/1 to: AAY15398 from: 1 to: 128  
12 CAGGTGCAGCTGCAGGAGTCAGGGGAGGAGTGTGTCAGGCTGGGGGCTC 61  
|||||  
1 GlnValGlnLeuGlnGluSerGlyGlyGlyLeuValGlnAlaGlyAspSe 17  
62 TCTGAGACTCTCCTGTGCAGCCTCGGGACGCGCCACGAGTGGTATGTC 111  
|||||  
17 rLeuArgLeuSerCysAlaAlaSerGlyArgThrSerHisGlyTyrGly 34  
112 ACTATGGTATGGCTGCTCCGCCAGGTTCCAGGAGACATGGTATAAAGACTCCGT 211  
|||||  
34 lyTyrGlyMetGlyTrpPheArgGlnIleProGlyLysGluArgGluLeu 50  
162 GTCGCAGCTATTAGGTGGAGTGGTAAAGAGACATGGTATAAAGACTCCGT 211  
|||||  
51 ValAlaAlaIleArgTrpSerGlyArgAsnThrTyrTyrAlaAspSerVa 67  
212 GAAGGGCCGATTACCATCTCCAGAGATAACGCCAAGACTACGGTTTATC 261  
|||||  
67 lLysGlyArgPheThrIleSerArgAspAsnValLysAspMetLeuTyrL 84  
262 TGCAAAATGAACAGCCTGAAACCTGAAGATACGGCGGTTTATTATTCGCC 311  
|||||  
84 euGlnMetAsnSerLeuLysProGluAspThrAlaValTyrThrCysAla 100  
312 GCTCGACCGTCCTCGGTGGATGATATTCCTGCGGTTGGGTTTACTA 361  
:::|||||  
101 ValArgThrValArgValValAspIleSerSerProValGlyPheAlaTy 117  
362 CTGGGGCCAGGGACCCAGGTCAACCGTCTCCTCA 395  
|||||  
117 rTrpGlyGlnGlyThrGlnValThrValSerSer 128

seq\_name: /SIDS5/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:AA15413

seq\_documentation\_block:  
ID AAY15413 standard; Protein; 250 AA.

XX  
AC AAY15413;  
XX  
DT 21-JUL-1999 (first entry)  
XX  
DE An anti-RR6 antigen binding protein.  
XX  
KW Multivalent antigen binding protein; single polypeptide chain;  
KW single domain binding site; diagnosis; therapy; targeting;  
KW immunoassay; cross-linking; agglutination; purification;  
KW phage inactivation; detergent.  
XX  
OS Synthetic.

XX  
PN WO9923221-A2.  
XX  
PD 14-MAY-1999.  
XX  
PF 27-OCT-1998; 98WO-EP06991.  
XX  
PR 27-OCT-1997; 97EP-0308538.  
XX  
PA (UNIL ) UNILEVER NV.  
PA (UNIL ) UNILEVER PLC.  
XX

PI Frenken LGJ, Howell S, Ledeboer AM, Van Der Logt CPE;  
XX WPI; 1999-313342/26.  
DR N-PSDB; AAX59616.

XX Multivalent antigen binding protein useful for inactivation of  
PT (bacterio)phages or viruses

XX Disclosure; Fig 12; 89pp; English.

XX The specification describes multivalent antigen binding proteins  
CC comprising a single polypeptide chain comprising, in series, two or  
CC more single domain binding sites. The multivalent antigen proteins  
CC are useful in applications where antibodies have been used in the  
CC prior art, including diagnosis, therapy, targeting, immunoassays,  
CC cross-linking methods including agglutination, or for purification  
CC processes. It is also useful for inactivation of (bacterio)phages  
CC or viruses, and in detergents. The present sequence represents an  
CC homodimeric bivalent anti-RR6 antigen binding protein.

XX Sequence 250 AA;

alignment\_scores:  
Quality: 550.50 Length: 161  
Ratio: 4.301 Gaps: 1  
Percent Similarity: 79.503 Percent Identity: 68.944

alignment\_block:  
US-09-742-690-1 x AAY15413 ..

Align seg 1/1 to: AAY15413 from: 1 to: 250

21 CTGAGGAGTCAGGGGAGGATGGTGCAGGCTGGGGCTCTCTGAGACT 70  
|||||  
1 LeuGlnGluSerGlyGlyLeuValGlnAlaGlyAspSerLeuArgLe 17

71 CTCCTGTCAGCCTCGGACGCGCCACCAAGTGGTCACTACTATGGTA 120  
|||||  
17 uSerCysAlaAlaSerGlyArgThrSerHisGlyTyrGlyTyrGlyM 34

121 TGGGCTGGTTCGCCAGGTTCCAGGGAAGGAGCGTGAGTTGTGCGAGCT 170  
|||||  
34 etGlyTrpPheArgGlnIleProGlyLysGluArgGluLeuValAlaAla 50

171 ATTAGTGGAGTGGTAAAGAGACATGGTATAAAGACTCCGTGAAGGGCCG 220  
|||||  
51 IleArgTrpSerGlyArgAsnThrTyrTyrAlaAspSerValLysGlyAr 67

221 ATTACCATCTCCAGAGATAACGCCAAGACTACGGTTTATCTGCAAAATGA 270  
|||||  
67 gPheThrIleSerArgAspAsnValLysAspMetLeuTyrLeuGlnMeta 84

271 ACAGCTGAACCTGAAGATACGCCCGTTTATTATTGTGCGCTCGACCG 320  
|||||  
84 snSerLeuLysProGluAspThrAlaValTyrThrCysAlaValArgThr 100

321 GTCCGGTGGATGATATTCCCTGCCGGTTGGGTTTGACTACTGGGGCCA 370  
|||||  
101 ValArgValValAspIleSerSerProValGlyPheAlaTyrTrpGlyGl 117

371 GGGGACCCAGGTCAACCGTCTCCTCAGGATCTCATCACCATCACCATCACG 420  
|||||  
117 nGlyThrGlnValThrValSerSerGlnValGlnLeuGlnGluSerGlyG 134

421 GA.....TCCACCTCCATTGAA 437  
|||  
134 lyGlyLeuValGlnAlaGlyAspSerLeuArgLeuSerCysAlaAlaSer 150

438 GGTCTGACCCAGTCTCAGTACGTCAGTGTGGT 470  
|||||  
151 GlyArgThrSerHisGlyTyrGlyTyrGly 161

seq\_name: /SID55/gcgdata/geneseq/geneseq-embl/AA2000.DAT:AA79104

seq\_documentation\_block:

ID AAY79104 standard; Protein; 128 AA.

XX AAY79104;

XX 23-MAY-2000 (first entry)

XX Llama heavy chain variable region recombinant fragment R7.

XX Antibody; heavy chain variable region; protein secretion;

XX glycosylation; llama; Saccharomyces cerevisiae.

XX Lama guanicoe glama.

XX WO200005389-A2.

XX 03-FEB-2000.

XX 08-JUL-1999; 99WO-EP04919.

XX 20-JUL-1998; 98EP-0202432.

XX (UNIL ) UNILEVER NV.

XX (UNIL ) UNILEVER PLC.

XX Frenken LGJ, Sagt C, Verkleij AJ, Verrips CT;

XX WPI; 2000-182710/16.

XX High level recombinant production of heterologous protein with  
PT increased secretory efficiency in lower eukaryotic cells, useful for  
PT producing lysozyme, cutinase, or antibodies -

XX Example 3d; Page 37; 55pp; English.

XX The present sequence is that of llama anti-Azo dye RR6 heavy chain  
CC variable region (VH) recombinant protein R7. Anti-RR6 VH DNA was  
CC amplified by PCR (see AA258489-91), ligated into yeast multicopy  
CC integration vectors and expressed in Saccharomyces cerevisiae.  
CC Anti-RR6 VH fragments R2, R5 and R7 were isolated and sequenced  
CC (see AAY79102-04). Secretion of the fragments was improved by the  
CC introduction at specific sites of N-glycosylation sites. The  
CC invention generally relates to the high level recombinant production  
CC of proteins, such as antibody fragments, in lower eukaryotic host  
CC (yeast or mould) cells. At least one N-glycosylation site is  
CC introduced into the protein at a location between the N-terminus  
CC and first hydrophobic region. The site is selected to increase the  
CC level of secretion without affecting the protein's functionality.

XX Sequence 128 AA;

alignment\_scores:

Quality: 544.00 Length: 128  
Ratio: 4.650 Gaps: 0  
Percent Similarity: 91.406 Percent Identity: 81.250

alignment\_block:

US-09-742-690-1 x AAY79104 ..

Align seg 1/1 to: AAY79104 from: 1 to: 128

12 CAGGTGCAGCTGCAGGAGTCAGGGGGAGGATGGTGCAGGCTGGGGGCTC 61

|||||

1 GlnValGlnLeuGlnGluSerGlyGlyGlyLeuValGlnThrGlyAspSe 17

62 TCTGAGACTCTCCTGTCAGCCCTCGGGACCGCCGCCACCATGTCATGGTC 111

|||||

17 rLeuArgLeuSerCysAlaAlaSerGlyArgThrSerHisGlyTyrGlyG 34

112 ACTATGGTATGGGCTGGTTCCGCCAGGTTCCAGGGGAAGGAGCGTGAGTTT 161



```
|||||
34 lyTyrGlyMetGlyTrpPheArgGlnIleProGlyLysGluArgGluLeu 50
|||||
162 GTCCGACCTATTAGTGGAGTGGTAAAGACACATGGTATAAAGACTCCGT 211
|||||
51 ValAlaAlaIleArgTrpSerGlyArgAsnThrTyrTyrAlaAspSerVa 67
|||||
212 GAAGGGCCGATTACCATCTCCAGAGATAACGCCAAGACTACGGTTTATC 261
|||||
67 lLysGlyArgPheThrIleSerArgAspAsnValLysAspMetLeuTyrL 84
|||||
262 TGCAATGAACAGCCTGAACCTGAAGATACGGCCGTTTATTTATGTGCC 311
|||||
84 euGlnMetAspSerLeuLysProGluAspThrAlaValThrTyrCysAla 100
|||||
312 GCTGACCGGTCCGCGTGATGATATTCCCTGCCGTTGGGTTTGACTA 361
:::
101 ValArgThrValArgValValAspIleSerSerProValGlyPheAlaTy 117
|||||
362 CTGGGGCCAGGGACCCAGGTCACCGTCTCCTCA 395
|||||
117 rTrpGlyGlnGlyThrGlnValThrValSerSer 128
|||||
```

seq\_name: /SIDS5/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:AA15406

seq\_documentation\_block:

ID AAY15406 standard; Protein; 238 AA.

XX AAY15406;

DT 21-JUL-1999 (first entry)

DE Aanti-hCG anti-RR6 antigen binding protein.

XX Multivalent antigen binding protein; single polypeptide chain;  
KW single domain binding site; diagnosis; therapy; targeting;  
KW immunoassay; cross-linking; agglutination; purification;  
KW phage inactivation; detergent.

OS Synthetic.

XX WO9923221-A2.

PD 14-MAY-1999.

XX 27-OCT-1998; 98WO-EP06991.

XX 27-OCT-1997; 97EP-0308538.

PA (UNIL ) UNILEVER NV.  
PA (UNIL ) UNILEVER PLC.

XX Frenken LGJ, Howell S, Ledebøer AM, Van Der Logt CPE;

XX WPI; 1999-313342/26.  
DR N-PSDB; AAX59612.

XX Multivalent antigen binding protein useful for inactivation of  
PT (bacterio)phages or viruses

XX Disclosure; Fig 27; 89pp; English.

XX The specification describes multivalent antigen binding proteins  
CC comprising a single polypeptide chain comprising, in series, two or  
CC more single domain binding sites. The multivalent antigen proteins  
CC are useful in applications where antibodies have been used in the  
CC prior art, including diagnosis, therapy, targeting, immunoassays,  
CC cross-linking methods including agglutination, or for purification  
CC processes. It is also useful for inactivation of (bacterio)phages  
CC or viruses, and in detergents. The present sequence represents an  
CC anti-hCG anti-RR6 bispecific biheaded antigen binding protein.

XX Sequence 238 AA;

alignment\_scores:  
Quality: 541.00 Length: 125  
Ratio: 4.746 Gaps: 0  
Percent Similarity: 91.200 Percent Identity: 82.400

alignment\_block:  
US-09-742-690-1 x AAY15406 ..

Align seg 1/1 to: AAY15406 from: 1 to: 238

```
12 CAGGTGCAGCTGCAGGAGTCAGGGGAGGATTGGTGCAGGCTGGGGGCTC 61
|||||
114 GlnValGlnLeuGlnGluSerGlyGlyGlyLeuValGlnAlaGlyAspSe 130
|||||
62 TCTGAGACTCTCCTGTGCAGCCTCGGACGCCACCAGTGGTCATGGTC 111
|||||
130 rLeuArgLeuSerCysAlaAlaSerGlyArgThrSerHisGlyTyrGlyG 147
|||||
112 ACTATGGTATGGGCTGGTTCGCCAGGTTCCAGGAGGAGGAGCGTGAGTTT 161
|||||
147 lyTyrGlyMetGlyTrpPheArgGlnIleProGlyLysGluArgGluLeu 163
|||||
162 GTCCAGCTATTAGTGGAGTGGTAAAGAGACATGGTATAAAGACTCCGT 211
|||||
164 ValAlaAlaIleArgTrpSerGlyArgAsnThrTyrTyrAlaAspSerVa 180
|||||
212 GAAGGGCCGATTACCATCTCCAGAGATAACGCCAAGACTACGGTTTATC 261
|||||
180 lLysGlyArgPheThrIleSerArgAspAsnValLysAspMetLeuTyrL 197
|||||
262 TGCAATGAACAGCCTGAACCTGAACATACGGCCGTTTATTTATGTGCC 311
|||||
197 euGlnMetAsnSerLeuLysProGluAspThrAlaValThrThrCysAla 213
|||||
312 GCTCGACCGGTCCGCGTGATGATATTTCCTGCGGTTGGGTTTGACTA 361
:::
214 ValArgThrValArgValValAspIleSerSerProValGlyPheAlaTy 230
|||||
362 CTGGGGCCAGGGACCCAGGTCAAC 386
|||||
230 rTrpGlyGlnGlyThrGlnValThr 238
|||||
```

seq\_name: /SIDS5/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:AA15408

seq\_documentation\_block:

ID AAY15408 standard; Protein; 239 AA.

XX AAY15408;

XX 21-JUL-1999 (first entry)

DE Anti-hCG anti-RR6 antigen binding protein.

XX Multivalent antigen binding protein; single polypeptide chain;  
KW single domain binding site; diagnosis; therapy; targeting;  
KW immunoassay; cross-linking; agglutination; purification;  
KW phage inactivation; detergent.

OS Synthetic.

XX WO9923221-A2.

PD 14-MAY-1999.

XX 27-OCT-1998; 98WO-EP06991.

XX 27-OCT-1997; 97EP-0308538.

PA (UNIL ) UNILEVER NV.  
PA (UNIL ) UNILEVER PLC.

XX

```
PI Frenken LGJ, Howell S, Ledeboer AM, Van Der Logt CPE;
XX
DR WPI; 1999-313342/26.
DR N-PSDB; AAX59614.
XX
PT Multivalent antigen binding protein useful for inactivation of
PT (bacterio)phages or viruses
XX
PS Disclosure; Fig 10; 89pp; English.
XX
CC The specification describes multivalent antigen binding proteins
CC comprising a single polypeptide chain comprising, in series, two or
CC more single domain binding sites. The multivalent antigen proteins
CC are useful in applications where antibodies have been used in the
CC prior art, including diagnosis, therapy, targeting, immunoassays,
CC cross-linking methods including agglutination, or for purification
CC processes. It is also useful for inactivation of (bacterio)phages
CC or viruses, and in detergents. The present sequence represents an
CC anti-hCG anti-RR6 bispecific biheaded antigen binding protein (ABP).
XX
SQ Sequence 239 AA;

alignment_scores:
  Quality: 541.00      Length: 125
  Ratio: 4.746        Gaps: 0
Percent Similarity: 91.200 Percent Identity: 82.400

alignment_block:
US-09-742-690-1 x AAY15408 ..

Align seg 1/1 to: AAY15408 from: 1 to: 239

12 CAGGTGCAGCTGCAGGAGTCAGGGGAGGATTGGTGCAGGCTGGGGCTC 61
   |||||
115 GlnValGlnLeuGlnGluSerGlyGlyGlyLeuValGlnAlaGlyAspSe 131
62 TCTGAGACTCTCCTGTGTGCAGCCTCGGGACGCGCCACCAGTGGTCATGGTC 111
   |||||
131 rLeuArgLeuSerCysAlaAlaSerGlyArgThrSerHisGlyTyrGlyG 148
112 ACTATGCTATGGCTGGTTCGCCAGGTTCCAGGGAAGGACGCTGAGTTT 161
   |||||
148 lyTyrGlyMetGlyTrpPheArgGlnIleProGlyLysGluArgGluLeu 164

162 GTCGCAGCTATTAGTGGAGTGGTAAAGAGACATGGTATAAGACTCCGT 211
   |||||
165 ValAlaAlaIleArgTrpSerGlyArgAsnThrTyrTyrAlaAspSerVa 181

212 GAAGGCGCGATTACCATCTCCAGAGATAACGCCAAGACTACGGTTTATC 261
   |||||
181 lLysGlyArgPheThrIleSerArgAspAsnValLysAspMetLeuTyrL 198

262 TGCAATGAACAGCCTGAACCTGAAGATACGGCCGTTTATTATTGTGCC 311
   |||||
198 euGlnMetAsnSerLeuLysProGluAspThrAlaValTyrThrCysAla 214

312 GCTCGACCGGTCCGCGTGGATGATATTTCCCTGCCGCTTGGGTTGACTA 361
   ::::|
215 ValArgThrValArgValValAspIleSerSerProvalGlyPheAlaTy 231

362 CTGGGCGCCAGGGACCCAGGTCACC 386
   |||||
231 rTrpGlyGlnGlyThrGlnValThr 239

seq_name: /SID55/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT:AAY15414
seq_documentation_block:
ID AAY15414 standard; Protein: 236 AA.
XX
AC AAY15414;
XX
DT 21-JUL-1999 (first entry)
```

```
XX An anti-RR6 antigen binding protein.
DE
XX
KW Multivalent antigen binding protein; single polypeptide chain;
KW single domain binding site; diagnosis; therapy; targeting;
KW immunoassay; cross-linking; agglutination; purification;
KW phage inactivation; detergent.
XX
OS Synthetic.
XX
XX WO9923221-A2.
PN
XX 14-MAY-1999.
PD
XX
XX 27-OCT-1998; 98WO-EP06991.
PF
XX 27-OCT-1997; 97EP-0308538.
PR
XX
XX (UNIL ) UNILEVER NV.
PA (UNIL ) UNILEVER PLC.
XX
XX Frenken LGJ, Howell S, Ledeboer AM, Van Der Logt CPE;
XX
DR WPI; 1999-313342/26.
DR N-PSDB; AAX59617.
XX
XX Multivalent antigen binding protein useful for inactivation of
PT (bacterio)phages or viruses
XX
PS Disclosure; Fig 13; 89pp; English.
XX
CC The specification describes multivalent antigen binding proteins
CC comprising a single polypeptide chain comprising, in series, two or
CC more single domain binding sites. The multivalent antigen proteins
CC are useful in applications where antibodies have been used in the
CC prior art, including diagnosis, therapy, targeting, immunoassays,
CC cross-linking methods including agglutination, or for purification
CC processes. It is also useful for inactivation of (bacterio)phages
CC or viruses, and in detergents. The present sequence represents an
CC heterodimeric bivalent anti-RR6 antigen binding protein.
XX
SQ Sequence 236 AA;

alignment_scores:
  Quality: 539.00      Length: 125
  Ratio: 4.728        Gaps: 0
Percent Similarity: 91.200 Percent Identity: 82.400

alignment_block:
US-09-742-690-1 x AAY15414 ..

Align seg 1/1 to: AAY15414 from: 1 to: 236

21 CTGCAGGAGTCAGGGGAGGATTGGTGCAGGCTGGGGCTCTCTGAGACT 70
   |||||
1 LeuGlnGluSerGlyGlyGlyLeuValGlnAlaGlyAspSerLeuArgLe 17

71 CTCTGTGCAGCTCGGGACGCGCCACCAGTGGTTCATGGTCACTATGGTA 120
   |||||
17 uSerCysAlaAlaSerGlyArgThrSerHisGlyTyrGlyGlyTyrGlyM 34

121 TGGCTGGTTCGCCAGGTTCCAGGGAAGGACGCTGAGTTTGTCCGAGCT 170
   |||||
34 etGlyTrpPheArgGlnIleProGlyLysGluArgGluLeuValAlaAla 50

171 ATTAGGTGGAGTGGTAAAGAGACATGGTATAAAGACTCCGTGAAGGCCG 220
   |||||
51 ileArgTrpSerGlyArgAsnThrTyrTyrAlaAspSerValLysGlyAr 67

221 ATTCACCATCTCCAGAGATAACGCCAAGACTACGGTTTATCTGCAATGA 270
   |||||
67 gPhethrIleSerArgAspAsnValLysAspMetLeuTyrLeuGlnMeta 84
```

271 ACAGCCTGAACCTGAAGATACGGCGTTATTATTGTGCCGCTCGACCG 320  
|||||  
84 snSerLeuLysProGluAspThrAlaValTyrThrCysAlaValArgThr 100  
|||||  
321 GTCCGCGTGGATGATATTCCCTGCGCGTGGGTTTGACTACTGGGGCCA 370  
|||||  
101 ValArgValValAspIleSerSerProValGlyPheAlaTyrTrpGlyG1 117  
|||||  
371 GGGGACCCAGGTACCGTCTCTCTCA 395  
|||||  
117 nGlyThrGlnValThrValSerSer 125  
|||||

seq\_name: /SIDS5/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:AA08986

seq\_documentation\_block:

ID AAY08986 standard; Protein; 147 AA.

XX AAY08986;

DT 20-AUG-1999 (first entry)

DE Llama protein construct pPIC.HCV3-His2t.

XX Analytical device; electrically conductive solid support; immobilisation;  
KW electroactive; analysis; peptide hormone; steroid; disease marker; HCV3;  
KW diagnostic indicator; antibody; bacteria; virus; fungi; yeast; llama.

OS Lama guanicoe glama.  
OS Synthetic.

XX WO9927356-A1.

XX 03-JUN-1999.

PF 23-NOV-1998; 98WO-GB03495.

PR 21-NOV-1997; 97EP-0309425.

XX (UNIL ) UNILEVER NV.  
PA (UNIL ) UNILEVER PLC.

PI Badley RA, Porter RA;

XX WPI; 1999-385228/32.

PT Analytical device including electrochemically active compound on  
PT solid support

XX Example 6.2; Page 67; 78pp; English.

CC This invention describes a novel component for an analytical device  
CC comprising an electrically conductive solid support on which a compound  
CC (I) is immobilized which has an electroactive part (EA) with an  
CC electrochemical property that can be modulated in a detectable manner by  
CC binding it to a specific binding partner (II). Compounds used in the  
CC method of the invention include those with the formula R = hydrogen,  
CC hydroxy, 1-14C alkyl, aryl, alkenyl or alkoxy (all optionally  
CC substituted), halo, amido, or amino; optionally one or more positions on  
CC the heteroaromatic rings are substituted by alkyl, aryl, alkenyl or  
CC alkoxy (all optionally substituted), acid groups (organic or inorganic),  
CC halo, amido or amino. Devices that contain the component of the invention  
CC are used for qualitative and quantitative analysis of e.g. nucleic acid,  
CC hormones (peptide or steroid), disease markers, diagnostic indicators  
CC etc., e.g. antibodies, bacteria, viruses, fungi or yeast. These devices  
CC can be used with turbid solutions (whole blood, milk, culture liquids  
CC etc.), unlike conventional colourimetric methods. This sequence  
CC represents a protein construct pPIC.HCV3-His2t which contains the llama  
CC HC-V fragment, HCV3 which is used in the method of the invention.

XX Sequence 147 AA;

alignment\_scores:  
Quality: 517.00 Length: 141  
Ratio: 4.419 Gaps: 2  
Percent Similarity: 82.979 Percent Identity: 73.050

alignment\_block:

US-09-742-690-1 x AAY08986 ..

Align seg 1/1 to: AAY08986 from: 1 to: 147

12 CAGGTGCAGCTGCAGGAGTCAGGGGAGGATGGTGCAGGCTGGGGCTC 61  
|||||  
1 GlnValGlnLeuGlnGluSerGlyGlyGlyLeuValGlnProGlyGlySe 17  
62 TCTGAGACTCTCCTGTGTGCAGCCTCGGGACGCCACACAGTGGTCATGGTC 111  
|||||  
17 rLeuArgLeuSerCysAlaAlaSerGlyLeuThrLeuThr.....T 31  
112 ACTATGGTATGGCTGGTTCGCCAGGTTCCAGGGAAGGAGCGTGAGTTT 161  
|||||  
31 hrTyrSerThrGlyTrpPheArgGlnAlaProGlyLysGluArgGluPhe 47  
162 GTCGCAGCTATTAGTGGAGT...GGTAAAGACATGGTATATAAGACTC 208  
|||||  
48 ValGlyMetLeuGlyTrpSerGlyGlyGlyAsnThrTyrTyrAlaAspSe 64  
209 CGTGAAGGGCCGATTACCATCTCCAGAGATACGCCAAGACTACGGTTT 258  
|||||  
64 rValLysGlyArgPheThrIleSerArgAspAsnAlaLysAsnMetValp 81  
259 ATCTGCAAAATGAACACCTGAAACCTGAAGATACGCCCGTTTATTATTGT 308  
:|||||  
81 heLeuGlnMetSerSerLeuLysProGluAspThrAlaValTyrTyrCys 97  
309 GCCGCTCGACCGGTCGCCGTGGATGATATTTCCTCGCGGTTGGTTGA 358  
|||||  
98 AlaAlaArgGlnProTyrArgGlySerTyrSerAspProAsnAsnTyrHi 114  
359 CTACTGGGGCCAGGGACCCAGGTCAACCGTCTCTCAGGATCTCATCACC 408  
|||||  
114 sTyrTrpGlyGlnGlyThrGlnValThrValSerSerGlySerHisHisH 131  
409 ATCACCATCACGGATCCACCTCC 431  
|||||  
131 iSHisHisGlySerGlySer 138

seq\_name: /SIDS5/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:AA08989

seq\_documentation\_block:

ID AAY08989 standard; Protein; 409 AA.

XX AAY08989;

DT 20-AUG-1999 (first entry)

DE Llama protein construct pPIC.scFv4155-link-HCV3-His2t.

XX Analytical device; electrically conductive solid support; immobilisation;  
KW electroactive; analysis; peptide hormone; steroid; disease marker; HCV3;  
KW diagnostic indicator; antibody; bacteria; virus; fungi; yeast; llama.

OS Lama guanicoe glama.

OS Synthetic.

XX WO9927356-A1.

XX 03-JUN-1999.

XX 23-NOV-1998; 98WO-GB03495.

XX 21-NOV-1997; 97EP-0309425.

XX (UNIL ) UNILEVER NV.

```
PA (UNIL ) UNILEVER PLC.
XX
PI Badley RA, Porter RA;
XX
DR WPI; 1999-385228/32.
XX
PT Analytical device including electrochemically active compound on
XX solid support
XX
PS Example 6.4; Page 70-71; 78pp; English.
XX
CC This invention describes a novel component for an analytical device
CC comprising an electrically conductive solid support on which a compound
CC (I) is immobilized which has an electroactive part (EA) with an
CC electrochemical property that can be modulated in a detectable manner by
CC binding it to a specific binding partner (II). Compounds used in the
CC method of the invention include those with the formula R = hydrogen,
CC hydroxy, 1-14C alkyl, aryl, alkenyl or alkoxy (all optionally
CC substituted), halo, amido, or amino; optionally one or more positions on
CC the heteroaromatic rings are substituted by alkyl, aryl, alkenyl or
CC alkoxy (all optionally substituted), acid groups (organic or inorganic),
CC halo, amido or amino. Devices that contain the component of the invention
CC are used for qualitative and quantitative analysis of e.g. nucleic acid,
CC hormones (peptide or steroid), disease markers, diagnostic indicators
CC etc., e.g. antibodies, bacteria, viruses, fungi or yeast. These devices
CC can be used with turbid solutions (whole blood, milk, culture liquids
CC etc.), unlike conventional colourimetric methods. This sequence
CC represents a protein construct pPIC.scFv4155-link-HCV3-His2t which
CC contains the llama HC-V fragment, HCV3 which is used in the method of
CC the invention.
XX
SQ Sequence 409 AA;

alignment_scores:
    Quality: 517.00      Length: 141
    Ratio: 4.419        Gaps: 2
    Percent Similarity: 82.979      Percent Identity: 73.050

alignment_block:
US-09-742-690-1 x AAY08989
..
Align seg 1/1 to: AAY08989 from: 1 to: 409

12 CAGTGCAGCTGTCAGGAGTCCAGGGGAGGATTGGTGCAGGCTGGGGGCTC 61
|||||
263 GlnValGlnLeuGlnGluSerGlyGlyGlyLeuValGlnProGlyGlySe 279

62 TCTGAGACTCTCCTGTGTGCAGCTCGGGACGCGCCACCAGTGGTCATGGTC 111
|||||
279 rLeuArgLeuSerCysAlaAlaSerGlyLeuThrLeuThr.....T 293

112 ACTATGGTATGGCTGGTTCGCCAGGTTCCAGGGAAGGAGCGTGAGTTT 161
|||||
293 hrTyrSerThrGlyTrpPheArgGlnAlaProGlyLysGluArgGluPhe 309

162 GTCGCAGCTATTAGGTGGAGT...GGTAAAGAGACATGGTATAAAGACTC 208
|||||
310 ValGlyMetLeuGlyTrpSerGlyGlyGlyAsnThrTyrTyrAlaAspSe 326

209 CGTGAAGGGCCGATTACCATCTCCAGAGATAACGCCAAGACTACGGTTT 258
|||||
326 rValLysGlyArgPheThrIleSerArgAspAsnAlaLysAsnMetValP 343

259 ATCTGCAAAATGAACAGCCTGAACCTGAAGATACGGCCGTTTATTATTGT 308
:::
343 heLeuGlnMetSerSerLeuLysProGluAspThrAlaValTyrTyrCys 359

309 GCCGCTCGACCGGTCCGGTGGATGATATTTCCTCGCCGGTGGGTTTGA 358
|||||
360 AlaAlaArgGlnProTyrArgGlySerTyrSerAspProAsnAsnTyrHi 376
:::
359 CTACTGGGGCCAGGGGACCCAGGTACCGTCTCCTCAGGATCTCATCACC 408
```

```
|||||
376 sTyrTrpGlyGlnGlyThrGlnValThrValSerSerGlySerHisHisH 393
|||||
409 ATCACCATCACGGATCCACCTCC 431
|||||
393 isHisHisHisGlySerGlySer 400

seq_name: /SID55/gcgcdata/geneseq/geneseq-emb1/AA1999.DAT: AAY08987

seq_documentation_block:
ID AAY08987 standard; Protein; 146 AA.
XX
AC AAY08987;
XX
DT 20-AUG-1999 (first entry)
XX
DE Llama protein construct pPIC.HCV24-His2t.
XX
KW Analytical device; electrically conductive solid support; immobilisation;
KW electroactive; analysis; peptide hormone; steroid; disease marker; HCV24;
KW diagnostic indicator; antibody; bacteria; virus; fungi; yeast; llama.
XX
OS Lama guanicoe glama.
OS Synthetic.
XX
PN WO9927356-A1.
XX
PD 03-JUN-1999.
XX
PF 23-NOV-1998; 98WO-GB03495.
XX
PR 21-NOV-1997; 97EP-0309425.
XX
PA (UNIL ) UNILEVER NV.
PA (UNIL ) UNILEVER PLC.
XX
PI Badley RA, Porter RA;
XX
DR WPI; 1999-385228/32.
XX
PT Analytical device including electrochemically active compound on
XX solid support
XX
PS Example 6.2; Page 68; 78pp; English.
XX
CC This invention describes a novel component for an analytical device
CC comprising an electrically conductive solid support on which a compound
CC (I) is immobilized which has an electroactive part (EA) with an
CC electrochemical property that can be modulated in a detectable manner by
CC binding it to a specific binding partner (II). Compounds used in the
CC method of the invention include those with the formula R = hydrogen,
CC hydroxy, 1-14C alkyl, aryl, alkenyl or alkoxy (all optionally
CC substituted), halo, amido, or amino; optionally one or more positions on
CC the heteroaromatic rings are substituted by alkyl, aryl, alkenyl or
CC alkoxy (all optionally substituted), acid groups (organic or inorganic),
CC halo, amido or amino. Devices that contain the component of the invention
CC are used for qualitative and quantitative analysis of e.g. nucleic acid,
CC hormones (peptide or steroid), disease markers, diagnostic indicators
CC etc., e.g. antibodies, bacteria, viruses, fungi or yeast. These devices
CC can be used with turbid solutions (whole blood, milk, culture liquids
CC etc.), unlike conventional colourimetric methods. This sequence
CC represents a protein construct pPIC.HCV24-His2t which contains the llama
CC HC-V fragment, HCV24 which is used in the method of the invention.
XX
SQ Sequence 146 AA;

alignment_scores:
    Quality: 508.50      Length: 140
    Ratio: 4.422        Gaps: 1
    Percent Similarity: 82.143      Percent Identity: 72.857

alignment_block:
```



```
US-09-742-690-1 x AAY08987  ..
Align seg 1/1  to: AAY08987  from: 1  to: 146

12 CAGGTGCAGCTGCAGGAGTCAGGGGAGGATTGGTCAGGCTGGGGGCTC 61
|||||
1  GlnValGlnLeuGlnGluSerGlyGlyGlyLeuValGlnAlaGlyGlySe 17
|||||
62 TCTGAGACTCTCCTGTGTCAGCCTCGGGACGCGCCACCAGTCGTGTCATGGTC 111
|||||
17 rLeuArgLeuSerCysAlaAlaSerGlyArgThrPheSer.....V 31
|||||
112 ACTATGGTATGGGCTGGTTCGCCAGGTTCCAGGGAAGAGCGTGAGTTT 161
|||||
31 aTyrAlaValGlyTrpPheArgGlnAlaProGlyLysGluArgGluPhe 47
|||||
162 GTCGACGCTATTAGTGGAGTGGTAAAGAGACATGGTATAAAGACTCCGT 211
|||||
48 ValGlyTyrPheGlyThrArgGlyGlyArgThrTyrTyrAlaAspSerVa 64
|||||
212 GAAGGCGCGATTACCATCTCCAGAGATAACGCCAAGACTACGGTTTATC 261
|||||
64 lLysGlyArgPheThrIleAlaIleAspAsnAlaLysAsnThrValTyrL 81
|||||
262 TGCAATGAACAGCCTGAAACCTGAAGATACGGCCGTTTATTGTGCC 311
|||||
81 euGlnMetAsnSerLeuLysLeuAspAspThrAlaValTyrTyrCysAla 97
|||||
312 GCTCGACCGGTCCGGTGGATGATATTTCCTGCCGGTTGGGTTTGACTA 361
|||||
98 ValArgMetProTyrSerGlyAspTyrArgSerSerGlyThrTyrAspTy 114
|||||
362 CTGGGGCCAGGGACCCAGGTCACCGTCTCCTCAGGATCTCATCACCATC 411
|||||
114 rTrpGlyGlnGlyThrGlnValThrValSerSerGlySerHisHisHish 131
|||||
412 ACCATCAGGATCCACCTCC 431
|||||
131 isHisHisGlySerGlySer 137
```

seq\_name: /SIDS5/gcgdata/geneseq/geneseq-emb1/AA1999.DAT: AAY08990

```
seq_documentation_block:
ID  AAY08990 standard; Protein; 408 AA.
XX
AC  AAY08990;
XX
DT  20-AUG-1999  (first entry)
XX
DE  Llama protein construct pPIC.scFv4155-link-HCV24-His2t.
XX
KW  Analytical device; electrically conductive solid support; immobilisation;
KW  electroactive; analysis; peptide hormone; steroid; disease marker; HCV24;
KW  diagnostic indicator; antibody; bacteria; virus; fungi; yeast; llama.
XX
OS  Lama guanicoe glama.
OS  Synthetic.
XX
PN  WO9927356-A1.
XX
PD  03-JUN-1999.
XX
PF  23-NOV-1998; 98WO-GB03495.
XX
PR  21-NOV-1997; 97EP-0309425.
XX
PA  (UNIL ) UNILEVER NV.
PA  (UNIL ) UNILEVER PLC.
XX
PI  Badley RA, Porter RA;
XX
DR  WPI; 1999-385228/32.
XX
```

```
PT Analytical device including electrochemically active compound on
XX solid support
PS Example 6.4; Page 72-73; 78pp; English.
XX
CC This invention describes a novel component for an analytical device
CC comprising an electrically conductive solid support on which a compound
CC (I) is immobilized which has an electroactive part (EA) with an
CC electrochemical property that can be modulated in a detectable manner by
CC binding it to a specific binding partner (II). Compounds used in the
CC method of the invention include those with the formula R - hydrogen,
CC hydroxy, 1-14C alkyl, aryl, alkenyl or alkoxy (all optionally
CC substituted), halo, amido, or amino; optionally one or more positions on
CC the heteroaromatic rings are substituted by alkyl, aryl, alkenyl or
CC alkoxy (all optionally substituted), acid groups (organic or inorganic),
CC halo, amido or amino. Devices that contain the component of the invention
CC are used for qualitative and quantitative analysis of e.g. nucleic acid,
CC hormones (peptide or steroid), disease markers, diagnostic indicators
CC etc., e.g. antibodies, bacteria, viruses, fungi or yeast. These devices
CC can be used with turbid solutions (whole blood, milk, culture liquids
CC etc.), unlike conventional colourimetric methods. This sequence
CC represents a protein construct pPIC.scFv4155-link-HCV24-His2t which
CC contains the llama HC-V fragment, HCV3 which is used in the method of
CC the invention.
XX
SQ Sequence 408 AA;
```

```
alignment_scores:
Quality: 508.50      Length: 140
Ratio: 4.422        Gaps: 1
Percent Similarity: 82.143  Percent Identity: 72.857

alignment_block:
US-09-742-690-1 x AAY08990  ..

Align seg 1/1  to: AAY08990  from: 1  to: 408
```

```
12 CAGGTGCAGCTGCAGGAGTCAGGGGAGGATTGGTCAGGCTGGGGGCTC 61
|||||
263 GlnValGlnLeuGlnGluSerGlyGlyGlyLeuValGlnAlaGlyGlySe 279
|||||
62 TCTGAGACTCTCCTGTGTCAGCCTCGGGACGCGCCACCAGTCGTGTCATGGTC 111
|||||
279 rLeuArgLeuSerCysAlaAlaSerGlyArgThrPheSer.....V 293
|||||
112 ACTATGGTATGGGCTGGTTCGCCAGGTTCCAGGGAAGAGCGTGAGTTT 161
|||||
293 aTyrAlaValGlyTrpPheArgGlnAlaProGlyLysGluArgGluPhe 309
|||||
162 GTCGACGCTATTAGTGGAGTGGTAAAGAGACATGGTATAAAGACTCCGT 211
|||||
310 ValGlyTyrPheGlyThrArgGlyGlyArgThrTyrTyrAlaAspSerVa 326
|||||
212 GAAGGCGCGATTACCATCTCCAGAGATAACGCCAAGACTACGGTTTATC 261
|||||
326 lLysGlyArgPheThrIleAlaIleAspAsnAlaLysAsnThrValTyrL 343
|||||
262 TGCAATGAACAGCCTGAAACCTGAAGATACGGCCGTTTATTGTGCC 311
|||||
343 euGlnMetAsnSerLeuLysLeuAspAspThrAlaValTyrTyrCysAla 359
|||||
312 GCTCGACCGGTCCGGTGGATGATATTTCCTGCCGGTTGGGTTTGACTA 361
|||||
360 ValArgMetProTyrSerGlyAspTyrArgSerSerGlyThrTyrAspTy 376
|||||
362 CTGGGGCCAGGGACCCAGGTCACCGTCTCCTCAGGATCTCATCACCATC 411
|||||
376 rTrpGlyGlnGlyThrGlnValThrValSerSerGlySerHisHisHish 393
|||||
412 ACCATCAGGATCCACCTCC 431
|||||
393 isHisHisGlySerGlySer 399
```